

THE DISTRIBUTION OF HOTELLING'S
GENERALIZED T_0^2

By
DAVID TIMOTHY HUGHES

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1970



*Dedicated to my wife Margaret and to my parents,
for their love and understanding*

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to Dr. J. G. Saw for his constant assistance and encouragement. Without his guidance, this work would not have been possible.

I would also like to thank Drs. P. V. Rao and Richard L. Scheaffer for their careful proofreading and helpful criticism. Special thanks must also go to Miss Gretchen Uzzell for her excellent editing and to Mrs. Carolyn Lyons who turned a crude manuscript into a masterpiece of typing.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vi
ABSTRACT	vii
Chapter	
I. STATEMENT OF THE PROBLEM	1
1.1 <i>Introduction</i>	1
1.2 <i>Historical Background</i>	6
1.3 <i>Summary of Results</i>	10
1.4 <i>Notation</i>	11
II. ORTHOSPHERICAL INTEGRALS	17
2.1 <i>Introduction</i>	17
2.2 <i>Some Relationships on the</i> <i>Orthospherical Integrals</i>	20
2.3 <i>Evaluation of Orthospherical</i> <i>Integrals</i>	28
III. MOMENTS OF HOTELLING'S GENERALIZED T_0^2	44
3.1 <i>Introduction</i>	44
3.2 <i>A Method for Evaluating the</i> <i>Moments of T_0^2</i>	46
3.3 <i>The First Two Moments of T</i>	56
3.4 <i>Evaluation of Certain Expectations</i> <i>Used in Obtaining the Moments</i> <i>of T</i>	64
3.5 <i>Summary</i>	73
IV. APPROXIMATIONS TO THE DISTRIBUTION OF HOTELLING'S GENERALIZED T_0^2	75
4.1 <i>Introduction</i>	75
4.2 <i>Approximations to the Distribution</i> <i>and Comparison of Percentage</i> <i>Points of T</i>	76
4.3 <i>Summary</i>	89

TABLE OF CONTENTS—*Continued*

Chapter	Page
V. THE POWER OF HOTELLING'S GENERALIZED T_0^2	92
5.1 <i>Introduction</i>	92
5.2 <i>An Upper Bound on the Integral of the Power Function of T_0^2</i>	94
5.3 <i>A Lower Bound on the Integral of the Power Function of T_0^2</i>	99
5.4 <i>The Monotonicity of the Power Function of T_0^2</i>	110
5.5 <i>Summary</i>	114
BIBLIOGRAPHY	117
BIOGRAPHICAL SKETCH	120

LIST OF TABLES

Table		Page
2.1	Orthospherical Integrals	39
4.1	Upper 5% points of T	83
4.2	Upper 1% points of T	84
4.3	Upper 5% points of T	85
4.4	Upper 1% points of T	86
4.5	Upper 5% points of T	87
4.6	Upper 1% points of T	88

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE DISTRIBUTION OF HOTELLING'S
GENERALIZED T_o^2

By

David Timothy Hughes

August, 1970

Chairman: Dr. J. G. Saw
Major Department: Statistics

Numerous criteria are available for testing the multivariate general linear hypothesis. In this work some properties of one of these criteria, T (a constant times Hotelling's generalized T_o^2), are investigated. The T statistic is the sum of the latent roots of a matrix obtained from random sampling from a multivariate normal population, and its distribution depends upon the corresponding population roots and some degrees of freedom parameters.

The orthospherical integrals involve expectations of the elements of a matrix having invariant measure on the orthogonal group. These integrals are defined and a method for evaluating them is presented. We have derived a special representation of T which, with the aid of the orthospherical integrals, enables us to present a method of evaluating the moments of T . The first two nonnull moments are obtained as an example.

The exact distribution of T in the general case is difficult to deal with and for this reason, three rather simple approximations to its distribution are obtained. All of these approximations are obtained by equating the moments of T with the moments of the approximating random variable. Accuracy comparisons of percentage points in certain special cases indicate the usefulness of two of these approximations.

A result by Wald and Hsu on the power function of the analysis of variance test is used to obtain an upper bound on the integral of the power function of T . A lower bound on the integral of the power function of T is obtained by using the properties of completely monotone functions. Completely monotone functions are also used to demonstrate the monotonicity of the power function of T .

CHAPTER I

STATEMENT OF THE PROBLEM

1.1 Introduction

Suppose we observe a number of m -variate column vectors, $\underline{x}_{ij}$: $i = 1, 2, \dots, k$; $j = 1, 2, \dots, n_i$ obtained by random sampling from each of k m -variate normal populations with common dispersion matrix, V . Assume that the elements of the column vectors associated with population i : $i = 1, 2, \dots, k$, have expectations given by the elements of an m -dimensional mean vector, $\underline{\mu}_i$. It is desired to test the hypothesis

$$H_0: \underline{\mu}_1 = \underline{\mu}_2 = \dots = \underline{\mu}_k$$

against

$$(1.1.1)$$

H_a : at least one of the above equalities does not hold.

With

$$N = \sum_{i=1}^k n_i, \quad (1.1.2)$$

define $\bar{\underline{x}}_{i.}$ and $\bar{\underline{x}}_{..}$ by

$$\bar{\underline{x}}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} \underline{x}_{ij}, \quad (1.1.3)$$

and

$$\bar{\underline{x}}_{..} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{n_i} \underline{x}_{ij}. \quad (1.1.4)$$

Construct the $m \times m$ matrix,

$$S_2 = \sum_{i=1}^k \sum_{j=1}^{n_i} (\underline{x}_{ij} - \bar{\underline{x}}_{i.}) (\underline{x}_{ij} - \bar{\underline{x}}_{i.})', \quad (1.1.5)$$

where $(\cdot)'$ denotes the transpose of the vector in brackets. The matrix S_2 will be nonsingular with probability one, whenever $N - k \geq m$ (see [1] or [32]). An appropriate test statistic for testing (1.1.1) is

$$\begin{aligned} T_0^2 &= (N - k) \sum_{i=1}^k n_i (\bar{\underline{x}}_{i.} - \bar{\underline{x}}_{..})' S_2^{-1} (\bar{\underline{x}}_{i.} - \bar{\underline{x}}_{..}) \\ &= (N - k) \operatorname{tr} S_1 S_2^{-1}, \end{aligned} \quad (1.1.6)$$

where S_1 is the $m \times m$ matrix defined by

$$S_1 = \sum_{i=1}^k n_i (\bar{\underline{x}}_{i.} - \bar{\underline{x}}_{..}) (\bar{\underline{x}}_{i.} - \bar{\underline{x}}_{..})'. \quad (1.1.7)$$

The null hypothesis is rejected if

$$T_0^2 > t, \quad (1.1.8)$$

where t is chosen to satisfy

$$P(T_0^2 > t | H_0) = \alpha \quad (1.1.9)$$

for some desired level of significance, α . The quantity, T_0^2 , defined in (1.1.6), has the distribution of Hotelling's generalized T_0^2 .

More generally, let $\underline{y}_i: i = 1, 2, \dots, v_1$, be m -dimensional, normally distributed column vectors with m -dimensional mean vectors, $\underline{\mu}_i$, and common $m \times m$ dispersion matrix, V . Suppose S_2 is an $m \times m$ central Wishart matrix (see [36]) distributed independently of $\underline{y}_1, \dots, \underline{y}_{v_1}$ with dispersion matrix V and degrees of freedom $v_2 > m - 1$. Then for testing the hypothesis

$$H_0: \underline{\mu}_1 = \dots = \underline{\mu}_{v_1} = \underline{0},$$

against (1.1.10)

$$H_a: \text{at least one } \underline{\mu}_i \neq \underline{0},$$

which is the canonical form of the general linear hypothesis, an appropriate test statistic is Hotelling's generalized T_0^2 , defined by

$$\begin{aligned} T_0^2 &= v_2 \sum_{i=1}^{v_1} \underline{y}_i' S_2^{-1} \underline{y}_i \\ &= v_2 \operatorname{tr} S_1 S_2^{-1}, \end{aligned} \quad (1.1.11)$$

where S_1 is the $m \times m$ matrix

$$S_1 = \sum_{i=1}^{v_1} \underline{y}_i \underline{y}_i'. \quad (1.1.12)$$

It is noted that S_1 defined in (1.1.12) has, in general, a noncentral Wishart distribution with dispersion matrix, V , degrees of freedom v_1 and $m \times m$ noncentrality matrix

$$MM' = \sum_{i=1}^{v_1} \underline{\mu}_i \underline{\mu}_i' . \quad (1.1.13)$$

In the first example given in which T_0^2 was defined by (1.1.6),

$$v_1 = k - 1, \quad (1.1.14)$$

$$v_2 = N - k, \quad (1.1.15)$$

and

$$MM' = \sum_{i=1}^k n_i (\underline{\mu}_i - \underline{\bar{\mu}}) (\underline{\mu}_i - \underline{\bar{\mu}})', \quad (1.1.16)$$

where

$$\underline{\bar{\mu}} = \frac{1}{N} \sum_{i=1}^k n_i \underline{\mu}_i . \quad (1.1.17)$$

The general problem just given is invariant under linear transformations. That is, if instead of observing $y_1, \dots, y_{v_1}$ and S_2 , we observe $\underline{z}_1, \dots, \underline{z}_{v_1}$ and S defined by

$$\underline{z}_i = A y_i : i = 1, \dots, v_1 \quad (1.1.18)$$

and

$$S = A S_2 A' \quad (1.1.19)$$

for any nonsingular $m \times m$ matrix, A , the problem remains the same. The m -variate column vectors, $\underline{z}_i: i = 1, \dots, v_1$, have mean vectors

$$\underline{\xi}_i = A\underline{\mu}_i : i = 1, \dots, v_1 \quad (1.1.20)$$

and the hypothesis (1.1.10) is equivalent to

$$H_0: \underline{\xi}_1 = \dots = \underline{\xi}_{v_1} = \underline{0},$$

against (1.1.21)

$$H_a: \text{at least one } \underline{\xi}_i \neq \underline{0}.$$

This property leads to consideration of test criteria which are also invariant under linear transformations. Any such criterion must be a function of the roots of the determinantal equation

$$|S_1 - \theta S_2| = 0, \quad (1.1.22)$$

where S_1 is defined in (1.1.12). The distribution of such a criterion is a function only of m, v_1, v_2 and the non-zero solutions of

$$|MM' - \lambda V| = 0. \quad (1.1.23)$$

(At most, $\min(m, v_1)$ of $\lambda_1, \dots, \lambda_m$ are nonzero.)

A number of test criteria have been proposed for testing (1.1.10). These include:

$$1. \text{ Likelihood ratio: } \prod_{i=1}^m (1 + \theta_i)^{-1}; \quad (1.1.24)$$

$$2. \text{ Roy's largest root: } \max(\theta_i); \quad (1.1.25)$$

$$3. \text{ Pillai's criteria: } \sum_{i=1}^m \theta_i (1 + \theta_i)^{-1}; \quad (1.1.26)$$

$$4. \text{ Hotelling's generalized } T_0^2: \nu_2 \sum_{i=1}^m \theta_i. \quad (1.1.27)$$

In this paper we investigate some of the properties of

$$T = T_0^2 / \nu_2 \quad (1.1.28)$$

for testing the general linear hypothesis. (It has become standard practice in recent years to study the properties of T rather than T_0^2 itself. We will continue in this practice and steadfastly hold that the symbol " T " will denote (1.1.28). The symbol " T_0^2 " will be used only when specifically discussing the properties of (1.1.11) rather than (1.1.28). The symbol " T " should not be confused with Hotelling's T^2 (see equation (1.2.1)). When discussing $(T_0^2 / \nu_2)^2$ we shall use " $(T)^2$ " and reserve " T^2 " for (1.2.1).

1.2 Historical Background

In 1951, Hotelling [11] suggested a statistic which he called T^2 as a generalization of "Student's t " ratio. Hotelling's T^2 is in fact a special case of T_0^2 given in (1.1.11) when $\nu_1 = 1$. Using (1.1.11), T^2 may be defined as

$$T^2 = v_2 \underline{Y}' S_2^{-1} \underline{Y}. \quad (1.2.1)$$

The distribution of T^2 was derived by Hotelling [11] under the null hypothesis, and, in 1938, Hsu [14] derived the nonnull distribution.

The statistic T (see (1.1.28)) was first suggested in 1938 by Lawley [26] as a generalization of Fisher's z (the familiar F ratio) for testing significance in multivariate analysis of variance. At that time, Lawley obtained the distribution of

$$z = \frac{1}{2} \log \left\{ \frac{v_2 - m - 1}{mv_2^2} T_0^2 \right\}, \quad (1.2.2)$$

and showed that the limiting distribution of T_0^2 as v_2 tends to infinity is the Chi-Square distribution on mv_1 degrees of freedom. Hsu [15] also demonstrated this result and obtained the null distribution of T in integral form. In this same paper, Hsu derived the first two nonnull moments of T , but, as we shall show, his second moment is in error.

Somewhat later, Hotelling [12] suggested T_0^2 as a generalization of his T^2 statistic for measuring multivariate dispersion. In 1951, the null distribution of T_0^2 was derived by Hotelling [13] for the case $m = 2$ and for general v_1 and $v_2 \geq 2$ in terms of the Gaussian hypergeometric function.

Since that time, numerous persons have studied the distribution of T_0^2 and T . Pillai [29] obtained the first three null moments of T and used these to obtain an approximation to its distribution. Using this approximation, Pillai and Samson [31] gave upper 5% and 1% points of T for $m = 2, 3$ and selected values of v_1 and v_2 . In 1956, Ito [17] obtained an asymptotic formula for the null distribution of T_0^2 as $v_2 \rightarrow \infty$ and in 1960 extended this result to cover the nonnull distribution [18]. Bagai [2] derived the null distribution of T_0^2 in terms of definite integrals for the cases $m = 2, 3$. The first two nonnull moments of T_0^2 were given by Mikhail [27] but his second moment, although it differs from that found by Hsu, is also incorrect.

The distributions of latent roots of random matrices were expressed by James [20], [22] and Constantine [3] in terms of zonal polynomials [3], [21]. Constantine [4] used these results to obtain the distribution of T , in both the null and nonnull case, in series form. These series converge, however, only for $|T| < 1$. In this same paper, Constantine derived the r th moment of T in terms of the zonal polynomials for $v_2 > 2r + m - 1$ (the moments do not exist otherwise) and listed the first two null moments explicitly.

In 1965, Khatri and Pillai [23] derived the first two moments of T in the linear case (the case of only one nonzero solution of (1.1.23)) and the same authors used

these results to approximate the distribution of T in the linear case [24]. They later generalized this to consider any number of nonzero roots [25]. Most recently, Davis [6] showed that, in the null case, the density of T satisfies an ordinary linear differential equation of degree m of Fuchsian type and obtains an equivalent first order system. These results are used to tabulate the upper 5% and 1% points of T_0^2/v_1 for $m = 3, 4$ [7]. Stein [34] showed that Hotelling's T^2 is an admissible test of

$$H_0: \underline{\mu} = \underline{0},$$

against

$$(1.2.3)$$

$$H_a: \underline{\mu} \neq \underline{0},$$

and Ghosh [10] extended this by showing that T_0^2 is admissible for testing the general linear hypothesis. Das Gupta, Anderson and Mudholkar [5] showed that the power of T_0^2 is an increasing function of each λ_i , the remaining λ_i 's being held fixed, where $\lambda_1, \dots, \lambda_m$ are defined in (1.1.23). Mikhail [27] and Srivastava [33] later claimed that the power of T_0^2 depends upon $\lambda_1, \dots, \lambda_m$ only through

$$\sum_{i=1}^m \lambda_i.$$

This claim is false, however, since the second moment of T_0^2 depends upon

$$\sum_{i=1}^m \lambda_i^2.$$

Ito [19] gave power comparisons of T_0^2 and the likelihood ratio criteria for the linear case and noted little or no difference in their powers for large v_2 . For the case $m = 2$, extensive power comparisons of T and three other test criteria for the general linear hypothesis have been made by Pillai and Jayachandran [30].

1.3 Summary of Results

Before proceeding with a discussion of the distribution of T , some preliminary mathematical results are needed. Therefore, Chapter II is devoted to the "orthospherical integrals." These integrals are defined and a method of evaluating them is presented. The orthospherical integrals up to order eight are tabulated at the end of Chapter II. In Chapter III, a general method for evaluating the moments of T is presented and the first two nonnull moments are found explicitly. These moments are used in Chapter IV to obtain three approximations to the distribution of T . The 5% and 1% points of T are found for each approximation, and accuracy comparisons are made with the exact percentage points and Pillai's approximate percentage points for $m = 2$. In the case $m = 3, 4$, comparisons are made with the percentage points given by Davis [7]. Upper and lower bounds are given for the integral of the power function of T . In addition we offer an alternative proof of the result due to Das Gupta, Anderson and Mudholkar.

1.4 Notation

Any work in multivariate analysis is greatly facilitated by frequent use of matrix notation, and it is also very convenient to abbreviate the distributional properties of random variables. For these reasons, certain notational conventions have been adopted. Although many of these are standard, they will be listed here for reference.

1. Matrices will be denoted by capital letters and the first time a matrix appears, its row and column dimensions will be given. Thus $A:(r \times s)$ denotes a matrix with r rows and s columns. The matrix having all its elements zero will be denoted by (0) . "I" will be reserved for the identity matrix.
2. The elements of a matrix will be denoted by the corresponding small letter with subscripts to denote their row and column position. Thus a_{ij} denotes the element in the i th row and j th column of the matrix A . The symbol $(A)_{ij}$ is equivalent to a_{ij} and is sometimes used for convenience.
3. An underscored small letter invariably represents a *column* vector. Its row dimension will be given the first time a vector appears. Thus $\underline{x}:(m)$ denotes a column vector consisting of m elements. The

vector which has all its elements zero will be denoted by $\underline{0}$.

4. It is sometimes convenient to form row or column vectors from the rows or columns of a matrix. The symbol $(A)_{i.}$ will denote the *row* vector formed from the i th row of the matrix A . The symbol $(A)_{.j}$ denotes the *column* vector formed from the j th column of the matrix A .
5. It is sometimes convenient to partition a square matrix $A:(r \times r)$ as

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right]$$

where A_{11} is the $p \times p$ matrix formed from the first p rows and columns A ; A_{12} is the $p \times q$ matrix formed from the first p rows and the last q columns of A , $p + q = r$. A_{21} and A_{22} are similarly defined.

6. The Kronecker product of $A:(r \times r)$ and $B:(s \times s)$ will be denoted by $A \otimes B$ and is a matrix $C:(rs \times rs)$ with

$$c_{(k-1)s+i, (l-1)s+j} = a_{kl} b_{ij}.$$

7. If x is a random variable having the normal density with mean μ , and variance σ^2 , this will be denoted by

$$x \sim N(u, \sigma^2).$$

8. If x is a random variable on $(0, \infty)$ with density

$$f(x) = \left\{ \Gamma\left(\frac{v}{2}\right) 2^{(v-2)/2} \right\}^{-1} x^{v-1} \exp\left\{-\frac{1}{2}x^2\right\},$$

we shall abbreviate this by

$$x \sim \chi_v(0).$$

This is to be read as "x has the central Chi density on v degrees of freedom."

9. If x is a random variable on $(0, \infty)$ with density

$$f(x) = \sum_{j=0}^{\infty} \exp\left(-\frac{\lambda}{2}\right) \left(\frac{\lambda}{2}\right)^j \left\{ j! \Gamma\left(\frac{v+2j}{2}\right) 2^{(v+2j)/2} \right\}^{-1} \\ \times x^{(v+2j)/2 - 1} \exp\left(-\frac{x}{2}\right),$$

we shall denote this by

$$x \sim \chi_v^2(\lambda).$$

This is to be read as "x has the noncentral Chi-square density on v degrees of freedom with non-centrality parameter, λ ." If $\lambda = 0$, the density becomes central.

10. If x_1 and x_2 are independently distributed with

$$x_1 \sim \chi_{v_1}^2(\lambda)$$

and

$$x_2 \sim \chi_{v_2}^2(0),$$

then

$$z = \frac{v_2}{v_1} \cdot \frac{x_1}{x_2}$$

has the noncentral F density on v_1 and v_2 degrees of freedom and noncentrality parameter, λ . This will be abbreviated

$$z \sim F_{v_1, v_2}(\lambda).$$

11. If x_1 and x_2 are independently distributed with

$$x_1 \sim \chi_{v_1}^2(\lambda)$$

and

$$x_2 \sim \chi_{v_2}^2(0),$$

then

$$u = \frac{x_2}{x_1 + x_2}$$

has the noncentral Beta density on v_1 and v_2 degrees of freedom with noncentrality parameter, λ . This will be denoted by

$$u \sim \beta e(v_1, v_2, \lambda).$$

12. If $\underline{x}$ is an m -dimensional column vector whose elements have a joint normal density with mean vector, $\underline{\mu}:(m)$ and dispersion matrix $V:(m \times m)$, this will be denoted by

$$\underline{x} \sim N_m(\underline{\mu}, V).$$

13. Let $\underline{x}_1, \dots, \underline{x}_n$ be independent m -variate column vectors with

$$\underline{x}_i \sim N_m(\underline{\mu}_i, V) : i = 1, \dots, n.$$

Then with $X = (\underline{x}_1, \dots, \underline{x}_n)$ and $M = (\underline{\mu}_1, \dots, \underline{\mu}_n)$, the joint density of the elements of X will be abbreviated

$$X \sim N_{m \times n}(M, V \otimes I).$$

14. With X an $m \times n$ matrix such that

$$X \sim N_{m \times n}(M, V \otimes I),$$

the $m \times m$ matrix

$$S = XX'$$

has a noncentral Wishart distribution with dispersion matrix V , degrees of freedom n , and noncentrality matrix MM' . This will be abbreviated

$$S \sim W_m(V, n, MM').$$

15. If S is an $m \times m$ symmetric matrix whose $m(m+1)/2$ mathematically independent elements have the density

$$f(S) = K_m(V, \nu) |S|^{(\nu-m-1)/2} \exp \left(-\frac{1}{2} \text{tr } V^{-1}S \right)$$

on S positive definite with

$$K_m(V, \nu) = \{ 2^{m\nu/2} \pi^{m(m-1)/4} |V|^{\nu/2} \prod_{j=1}^m \Gamma\left(\frac{\nu+1-j}{2}\right) \}^{-1}$$

this will be abbreviated by

$$S \sim W_m(V, v, (0)).$$

This is to be read as "the $m(m + 1)/2$ mathematically independent elements of S have the central Wishart density with dispersion matrix, V , and degrees of freedom v ."

16. The symbols " $a \implies b$ " and " $a \Longleftarrow b$ " are to be read " a implies b " and " a implies, and is implied by, b ," respectively.

17. Suppose it is desired to integrate the function $f(x_1, \dots, x_n)$ over the set

$$A = \{(x_1, \dots, x_n) : a_i \leq x_i \leq b_i : i = 1, \dots, n\}.$$

Then

$$\int_A f(\underline{x}) d(\underline{x})$$

will be used to abbreviate

$$\int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \dots dx_1.$$

18. For referencing within the text $[\cdot]$ will denote bibliographical references, while $(\cdot)$ will denote references to equations. Thus $[8]$ refers to the eighth entry in the bibliography, while $(4.2.6)$ refers to equation six of Section 2 in Chapter IV.

Other special notational conventions will be introduced as they occur in the text.

CHAPTER II

ORTHOSPHERICAL INTEGRALS

2.1 Introduction

Let $\Omega(m \times m)$ denote the group of $m \times m$ orthogonal matrices. Let Φ be an $m \times m$ matrix of random variables with Φ belonging to the group $\Omega(m \times m)$, and suppose that Φ has invariant measure on $\Omega(m \times m)$. This measure can be determined in the following manner. Consider any row of Φ , say the first row for convenience. Then since Φ is orthogonal, the elements, ϕ_{1j} : $j = 1, \dots, m$, of $(\Phi)_{1.}$ must satisfy

$$\begin{aligned} (\Phi)_{1.} (\Phi)_{1.}' &= \sum_{j=1}^m \phi_{1j}^2 \\ &= 1. \end{aligned} \tag{2.1.1}$$

This is equivalent to saying that $(\Phi)_{1.}$ determines a point on an m -dimensional hypersphere of unit radius. If Φ is to have invariant measure on $\Omega(m \times m)$, then the measure on $(\Phi)_{1.}$ must be

$$\begin{aligned} f[(\Phi)_{1.}] &= \frac{1}{S_m(1)} \quad \text{if } (\Phi)_{1.} (\Phi)_{1.}' = 1; \\ &= 0 \quad \text{otherwise,} \end{aligned} \tag{2.1.2}$$

where

$$S_j(r) = 2\pi^{j/2} r^{j-1} / \Gamma(j/2) \tag{2.1.3}$$

is the surface area of a hypersphere of radius r in j dimensions.

Now fix the elements in the first row of Φ , and consider any other row, say the second row for convenience. Due to the orthogonality of Φ , the elements, ϕ_{2j} : $j = 1, \dots, m$, of $(\Phi)_2$. must satisfy

$$\sum_{j=1}^m \phi_{1j} \phi_{2j} = 0,$$

and (2.1.4)

$$\sum_{j=1}^m \phi_{2j}^2 = 1.$$

This implies that $(\Phi)_2$. determines a point lying on an $(m - 1)$ -dimensional hypersphere of unit radius which is orthogonal to the vector $(\Phi)_1$.. Since Φ is to have invariant measure on $\Omega(m \times m)$, the measure on $(\Phi)_2$. conditional on $(\Phi)_1$. must be

$$\begin{aligned} f[(\Phi)_2. | (\Phi)_1.] &= \frac{1}{S_{m-1}(1)} \quad \text{if (2.1.4) is satisfied;} \\ &= 0 \quad \text{otherwise.} \end{aligned} \tag{2.1.5}$$

Continuing in this manner, fix the elements of $(\Phi)_1., (\Phi)_2., \dots, (\Phi)_{i-1},.$ and consider the i th row, $(\Phi)_i$., of Φ . Since Φ is to be orthogonal, the elements, ϕ_{ij} : $j = 1, \dots, m$, of $(\Phi)_i$. must satisfy

$$\sum_{j=1}^m \phi_{1j} \phi_{ij} = 0,$$

$$\sum_{j=1}^m \phi_{2j} \phi_{ij} = 0,$$

·
·
·

(2.1.6)

$$\sum_{j=1}^m \phi_{i-1,j} \phi_{ij} = 0,$$

and

$$\sum_{j=1}^m \phi_{ij}^2 = 1.$$

This implies that $(\Phi)_{i.}$ determines a point on an $(m-i+1)$ -dimensional hypersphere of unit radius which is orthogonal to the mutually orthogonal vectors $(\Phi)_{1.}, (\Phi)_{2.}, \dots, (\Phi)_{i-1.}$. Since Φ is to have invariant measure on $\Omega(m \times m)$, the measure on $(\Phi)_{i.}$, conditional on $(\Phi)_{1.}, (\Phi)_{2.}, \dots, (\Phi)_{i-1.}$, is

$$f \left[(\Phi)_{i.} \mid (\Phi)_{1.}, (\Phi)_{2.}, \dots, (\Phi)_{i-1.} \right] = \frac{1}{S_{m-i+1}(1)}$$

if (2.1.6) is satisfied; (2.1.7)

= 0 otherwise.

Hence the measure on Φ is

$$f(\Phi) = \prod_{j=1}^m \frac{1}{S_j(1)} \quad \text{if } \Phi\Phi' = I;$$

(2.1.8)

= 0 otherwise.

Using (2.1.2), we can write the measure on Φ as

$$\begin{aligned} f(\Phi) &= \prod_{j=1}^m \frac{\Gamma(j/2)}{2\pi^{j/2}} \quad \text{if } \Phi\Phi' = I; \\ &= 0 \quad \text{otherwise.} \end{aligned} \quad (2.1.9)$$

In the derivation of the moments of T , we shall need certain moments and product moments of the elements of Φ . To this end, for any $m \times m$ matrix, C , of nonnegative integers, define

$$\begin{aligned} J(C) &= E \left[\prod_{i=1}^m \prod_{j=1}^m \phi_{ij}^{c_{ij}} \right] \\ &= \int_{\Omega(m \times m)} \left[\prod_{i=1}^m \prod_{j=1}^m \phi_{ij}^{c_{ij}} \right] f(\Phi) d\Phi. \end{aligned} \quad (2.1.10)$$

These are the "orthospherical integrals." We shall say that $J(C)$ is of order s , where

$$s = \sum_{i=1}^m \sum_{j=1}^m c_{ij}. \quad (2.1.11)$$

2.2 Some Relationships on the Orthospherical Integrals

Let P and Q be fixed $m \times m$ orthogonal matrices and define an $m \times m$ random matrix, Θ , by

$$\Theta = P\Phi Q, \quad (2.2.1)$$

wherein Φ is random with invariant measure on $\Omega(m \times m)$. Then $\Theta \in \Omega(m \times m)$ and, since it is obtained from Φ by an orthogonal transformation, Θ also has invariant measure on $\Omega(m \times m)$.

Now

$$\begin{aligned} J(C) &= E \left[\prod_{i=1}^m \prod_{j=1}^m \theta_{ij}^{c_{ij}} \right] \\ &= E \left[\prod_{i=1}^m \prod_{j=1}^m \left(\sum_{\alpha=1}^m \sum_{\beta=1}^m p_{i\alpha} \phi_{\alpha\beta} q_{\beta j} \right)^{c_{ij}} \right]. \end{aligned} \quad (2.2.2)$$

By suitable choice of P and Q , (2.2.2) yields a number of interesting relationships on the orthospherical integrals.

In particular, suppose that P is obtained from the $m \times m$ identity matrix, I , by some permutation of the rows or columns of I , and that Q is similarly obtained by a possibly different permutation. P and Q are therefore orthogonal, and clearly they have one and only one element in each row and column equal to unity and all other elements zero. That is,

$$\begin{aligned} p_{i\alpha} &= 1 & \text{if } \alpha &= i_i^* & \text{for some } i_i^*, \\ &= 0 & \text{if } \alpha &\neq i_i^*, \end{aligned} \quad (2.2.3)$$

where i_i^* differs for each i , and

$$\begin{aligned} q_{\beta j} &= 1 & \text{if } \beta &= j_j^* & \text{for some } j_j^*, \\ &= 0 & \text{if } \beta &\neq j_j^*, \end{aligned} \quad (2.2.4)$$

where j_j^* differs for each j . Now

$$\sum_{\alpha=1}^m \sum_{\beta=1}^m p_{i\alpha} \phi_{\alpha\beta} q_{\beta j} = \phi_{i i^*, j j^*}, \quad (2.2.5)$$

so that (2.2.2) can be written as

$$J(C) = E \left[\prod_{i=1}^m \prod_{j=1}^m \phi_{i i^*, j j^*}^{c_{ij}} \right], \quad (2.2.6)$$

with $\phi_{i i^*, j j^*}$ the elements of Θ , where Θ is defined in (2.2.1). We may rewrite (2.2.6) as

$$J(C) = E \left[\prod_{i=1}^m \prod_{j=1}^m \phi_{ij}^{c_{i i^*, j j^*}} \right], \quad (2.2.7)$$

with $c_{i i^*, j j^*}$ the elements of $C^* : (m \times m)$ defined by

$$C^* = P' C Q'. \quad (2.2.8)$$

That is, C^* is obtained from C by a permutation of the rows or columns of C . Since by definition

$$J(C^*) = E \left[\prod_{i=1}^m \prod_{j=1}^m \phi_{ij}^{c_{i i^*, j j^*}} \right], \quad (2.2.9)$$

then

$$J(C) = J(C^*). \quad (2.2.10)$$

The permutation matrices, P and Q , were obtained arbitrarily so that (2.2.10) holds for all permutations of C . Thus, for example, with $m = 4$

$$\begin{aligned}
 J \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} &= J \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 &= J \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 &= J \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix},
 \end{aligned}$$

and so on.

Equation (2.2.8) can be used to simplify the notation by displaying only the smallest rectangular array containing all of the nonzero elements of C , along with an index, m , to denote the dimensions of Φ and C . If C is the smallest rectangular array containing all nonzero elements of C , then

$$J_m(C) = J(C). \quad (2.2.11)$$

For example, with $m = 4$,

$$J_4 \begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = J \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= J \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and so on.

Let P and Q be diagonal matrices whose diagonal elements are either ± 1 , so that P and Q are orthogonal.

Then from (2.2.2) and (2.2.11),

$$\begin{aligned} J_m(C) &= E \left[\prod_{i=1}^m \prod_{j=1}^m (p_{ii} q_{jj})^{c_{ij}} \phi_{ij}^{c_{ij}} \right] \\ &= J_m(C) \prod_{i=1}^m \prod_{j=1}^m (p_{ii} q_{jj})^{c_{ij}} \\ &= J_m(C) \prod_{i=1}^m p_{ii}^{c_{i.}} \prod_{j=1}^m q_{jj}^{c_{.j}}, \end{aligned} \quad (2.2.12)$$

where

$$c_{i.} = \sum_{j=1}^m c_{ij},$$

and (2.2.13)

$$c_{.j} = \sum_{i=1}^m c_{ij}.$$

Suppose that $c_{i.}$ is an odd integer for some $i \in \{1, \dots, m\}$.

For convenience, assume that $c_{1.}$ is odd. Setting

$$\begin{aligned} p_{11} &= -1, \\ p_{ii} &= +1: & i &= 2, \dots, m, \\ q_{jj} &= +1: & j &= 1, \dots, m, \end{aligned} \quad (2.2.14)$$

equation (2.2.12) gives

$$J_m(C) = -J_m(C). \quad (2.2.15)$$

It follows that

$$J_m(C) = 0 \quad (2.2.16)$$

whenever any row or column of C sums to an odd integer.

Thus, for example,

$$J_m \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix} = 0,$$

and

$$J_m \begin{pmatrix} 4 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 0.$$

By considering other special matrices, P and Q , many other relationships can be obtained on the orthospherical integrals. For example, choosing

$$P = \left[\begin{array}{cc|c} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & (0) \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \\ \hline (0) & & I \end{array} \right],$$

and

$$(2.2.17)$$

$$Q = I,$$

equation (2.2.1) yields

$$\theta_{11} = \frac{1}{\sqrt{2}} (\phi_{11} + \phi_{21})$$

and (2.2.18)

$$\theta_{22} = \frac{1}{\sqrt{2}} (\phi_{12} - \phi_{22}).$$

Hence

$$\begin{aligned} J_m \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} &= E(\theta_{11}^2 \theta_{22}^2) \\ &= \frac{1}{4} E(\phi_{11} + \phi_{21})^2 (\phi_{12} - \phi_{22})^2. \end{aligned} \quad (2.2.19)$$

Expanding (2.2.19) and making use of (2.2.16) gives

$$2 J_m \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = J_m(2,2) - J_m \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \quad (2.2.20)$$

Thus $J_m \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ can be evaluated once $J_m(2,2)$ and $J_m \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ are known.

Other relationships can be obtained by simply making use of the orthogonality of Φ . Some examples are presented below.

$$\begin{aligned} 1. \quad \phi_{11}^2 &= \phi_{11}^2 \left(\sum_{j=1}^m \phi_{1j}^2 \right) \\ &= \phi_{11}^4 + \sum_{j=2}^m \phi_{11}^2 \phi_{1j}^2. \end{aligned} \quad (2.2.21)$$

Taking expectations in (2.2.21) yields

$$J_m(2) = J_m(4) + (m - 1) J_m(2, 2). \quad (2.2.22)$$

$$\begin{aligned} 2. \quad \phi_{11}^2 \phi_{12}^2 &= \phi_{11}^2 \phi_{12}^2 \left(\sum_{i=1}^m \phi_{i2}^2 \right) \\ &= \phi_{11}^2 \phi_{12}^4 + \sum_{i=2}^m \phi_{11}^2 \phi_{12}^2 \phi_{i2}^2. \end{aligned} \quad (2.2.23)$$

Taking expectations yields

$$J_m(2, 2) = J_m(2, 4) + (m - 1) J_m \left(\begin{smallmatrix} 2 & 2 \\ 0 & 2 \end{smallmatrix} \right). \quad (2.2.24)$$

$$\begin{aligned} 3. \quad 0 &= \sum_{j=1}^m \phi_{1j} \phi_{2j} \\ &= \phi_{11} \phi_{21} \left(\sum_{j=1}^m \phi_{1j} \phi_{2j} \right) \\ &= \phi_{11}^2 \phi_{21}^2 + \sum_{j=2}^m \phi_{11} \phi_{21} \phi_{1j} \phi_{2j}. \end{aligned} \quad (2.2.25)$$

Thus

$$J_m(2, 2) + (m - 1) J_m \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix} \right) = 0. \quad (2.2.26)$$

Such relationships greatly simplify the evaluation of the integrals, since they reduce the number of integrals which must be evaluated directly. Direct evaluation of the orthospherical integrals can be tedious, but a relatively easy method is presented in the next section.

2.3 Evaluation of Orthospherical Integrals

The orthospherical integrals which need to be evaluated directly may be evaluated by a number of methods. A polar transformation on the elements of Φ may be used when the integrals involve only the elements in a single row or column of Φ . This method becomes extremely cumbersome, however, when elements of more than one row or column are involved. Perhaps the simplest method is one which depends upon the following representation of Φ .

Let X be an $m \times m$ matrix of random variables and suppose

$$X \sim N_{m \times m}((0), I \otimes I). \quad (2.3.1)$$

Define an $m \times m$ matrix R to be a lower triangular matrix satisfying

$$XX' = RR'. \quad (2.3.2)$$

Equation (2.3.2) can be written as a system of equations on the element of X and R .

$$\sum_{j=1}^m x_{1j}^2 = r_{11}^2,$$

$$\sum_{j=1}^m x_{1j}x_{ij} = r_{11}r_{i1}: \quad i = 2, \dots, m,$$

$$\sum_{j=1}^m x_{2j}^2 = r_{21}^2 + r_{22}^2.$$

$$\sum_{j=1}^m x_{2j} x_{ij} = r_{21} r_{i1} + r_{22} r_{i2}: \quad i = 3, \dots, m,$$

.
.
.

and

$$\sum_{j=1}^m x_{mj}^2 = \sum_{j=1}^m r_{mj}^2. \quad (2.3.3)$$

If the diagonal elements, $r_{ii}: i = 1, \dots, m$, are required to be nonnegative, then the equations (2.3.3) uniquely determine the elements of R . Hence, R is the unique lower triangular $m \times m$ matrix with nonnegative diagonal elements satisfying (2.3.2). We shall call R the lower triangular square root of XX' .

Define a matrix $\Phi: (m \times m)$ by

$$X = R\Phi. \quad (2.3.4)$$

It is to be noted that the elements of X , and hence of R , are continuous random variables. Therefore, the diagonal elements, $r_{ii}: i = 1, \dots, m$, will be positive with probability one, and hence R will be nonsingular with probability one. Then Φ is properly defined by (2.3.4) and

$$\Phi = R^{-1}X. \quad (2.3.5)$$

Using (2.3.2) and (2.3.5), Φ must satisfy

$$\Phi\Phi' = (R^{-1}X)(X'R'^{-1})$$

$$\begin{aligned}
 &= R^{-1}(XX')R'^{-1} \\
 &= R^{-1}(RR')R'^{-1} \\
 &= I,
 \end{aligned}
 \tag{2.3.6}$$

so that ϕ has zero measure except on a subset of $\Omega(m \times m)$. Let $f(X)$ denote the joint density of the elements of X described by (2.3.1), and P be any fixed, orthogonal, $(m \times m)$ matrix. Since $f(X) > 0$ for any real X , then $f(XP) > 0$. From (2.3.2), and the fact that P is orthogonal, we have

$$\begin{aligned}
 RR' &= XX' \\
 &= XPP'X' \\
 &= (XP)(XP)'.
 \end{aligned}
 \tag{2.3.7}$$

Thus the lower triangular square root, R , of XX' is identical to that of $XPP'X'$. Then

$$X = R\phi \implies XP = R\phi P. \tag{2.3.8}$$

This implies that whatever ϕ has nonzero measure, then so does ϕP . If $g(\phi)$ is the measure on ϕ then

$$g(\phi) > 0 \quad \text{for all } \phi \in \Omega(m \times m). \tag{2.3.9}$$

The matrix ϕ will be shown to have invariant measure on $\Omega(m \times m)$. The joint density of ϕ and R will be obtained, and it will be shown that the elements of R are both mutually independent and independent of the elements

of ϕ . The joint density of R and ϕ will be denoted by $h(R, \phi)$. This joint density can be obtained using the density of X given by (2.3.1), and obtaining the Jacobian of the transformation from X to R and ϕ defined by (2.3.4) and (2.3.6). Denoting this Jacobian by $J(X \rightarrow R, \phi)$, then

$$h(R, \phi) = f(R\phi) J(X \rightarrow R, \phi), \quad (2.3.10)$$

where $f(\cdot)$ is the density of (2.3.1). Direct evaluation of the Jacobian is cumbersome and the method used here is due to Hsu and reported by Deemer and Olkin [8].

We seek the Jacobian of the transformation from X to R and ϕ defined by

$$X = R\phi$$

$$\text{and} \quad (2.3.11)$$

$$\phi\phi' = I,$$

with all matrices $m \times m$ and R lower triangular. Let δX , δR , and $\delta\phi$ be $m \times m$ matrices of elements, $(\delta X)_{ij}$, $(\delta R)_{ij}$, and $(\delta\phi)_{ij}$, denoting small changes in the elements of X , R , and ϕ respectively. Suppose that the changes δR in R and $\delta\phi$ in ϕ bring about a change δX in X so that (2.3.11) is preserved. That is,

$$X + \delta X = (R + \delta R)(\phi + \delta\phi)$$

$$\text{and} \quad (2.3.12)$$

$$(\phi + \delta\phi)(\phi + \delta\phi)' = I.$$

Expanding equations (2.3.12) and dropping terms of second order in the δ 's, we find that

$$X + \delta X = R\Phi + R(\delta\Phi) + (\delta R) \Phi$$

and (2.3.13)

$$\Phi\Phi' + \Phi(\delta\Phi)' + (\delta\Phi)\Phi' = I.$$

Making use of (2.3.11), we can write equations (2.3.13)

as

$$\delta X = R(\delta\Phi) + (\delta R) \Phi$$

and (2.3.14)

$$\Phi(\delta\Phi)' + (\delta\Phi) \Phi' = 0.$$

Hsu has shown that

$$J(X \rightarrow R, \Phi) = J(\delta X \rightarrow \delta R, \delta\Phi), \quad (2.3.15)$$

where $J(\delta X \rightarrow \delta R, \delta\Phi)$ is the Jacobian of the transformation from δX to δR and $\delta\Phi$ defined by equations (2.3.14), in which R and Φ are now considered to be fixed $m \times m$ matrices.

For convenience, denote δX by A , δR by B , and $\delta\Phi$ by C . Then

$$J(X \rightarrow R, \Phi) = J(A \rightarrow B, C), \quad (2.3.16)$$

with

$$A = RC + B\Phi$$

and (2.3.17)

$$C\Phi' + \Phi C' = 0.$$

Define $m \times m$ matrices A^* and C^* by

$$A^* = A\Phi'$$

and

$$(2.3.18)$$

$$C^* = C\Phi'.$$

Using (2.3.17) and (2.3.18),

$$A^* = RC^* + B$$

and

$$(2.3.19)$$

$$C^* + C^{*'} = 0,$$

so that C^* is an asymmetric matrix. Furthermore,

$$J(A \rightarrow B, C) = J(A \rightarrow A^*) J(A^* \rightarrow B, C^*) J(C^* \rightarrow C). \quad (2.3.20)$$

The Jacobians, $J(A \rightarrow A^*)$ and $J(C^* \rightarrow C)$, are the Jacobians of the transformations defined by (2.3.18). These are orthogonal transformations and hence,

$$J(A \rightarrow A^*) = J(C^* \rightarrow C) = 1. \quad (2.3.21)$$

The Jacobian $J(A^* \rightarrow B, C^*)$ is the Jacobian of the transformation defined by (2.3.19). Recall that $B = \delta R$, and hence is lower triangular. Direct evaluation of this Jacobian can be carried out rather easily. It is in fact

$$J(A^* \rightarrow B, C^*)$$

$$= \left\| \frac{\partial (a_{11}^* a_{21}^* a_{22}^* \cdots a_{m1}^* a_{m2}^* \cdots a_{mm}^* a_{12}^* a_{13}^* \cdots a_{1m}^* a_{23}^* \cdots a_{m-1,m}^*)}{\partial (b_{11}^* b_{21}^* b_{22}^* \cdots b_{m1}^* b_{m2}^* \cdots b_{mm}^* c_{12}^* c_{13}^* \cdots c_{1m}^* c_{23}^* \cdots c_{2m}^* \cdots c_{m-1,m}^*)} \right\|. \quad (2.3.22)$$

Laying out this matrix of derivatives in tabular form, we have

$a_{11}^* a_{21}^* a_{22}^* \dots a_{m1}^* a_{m2}^* \dots a_{mm}^*$	$a_{12}^* a_{13}^* \dots a_{1m}^* a_{23}^* \dots a_{2m}^* \dots a_{m-1,m}^*$
b_{11}	(0)
b_{21}	
b_{22}	
.	
.	
b_{m1}	
b_{m2}	
.	
.	
.	
b_{mm}	
c_{12}^*	$r_{11} \quad 0 \quad \dots 0 \quad 0 \quad \dots 0 \quad \dots 0$
c_{13}^*	$0 \quad r_{11} \quad \dots 0 \quad r_{21} \quad \dots 0 \quad \dots 0$
.	.
.	.
.	.
c_{1m}^*	$0 \quad 0 \quad \dots r_{11} \quad 0 \quad \dots r_{21} \quad \dots r_{m-1,1}$
c_{23}^*	$0 \quad 0 \quad \dots 0 \quad r_{22} \quad \dots \quad \dots$
.	.
.	.
.	.
c_{2m}^*	$\vdots \quad 0 \quad \dots 0 \quad 0 \quad \dots r_{22} \quad \dots r_{m-1,2}$
.	.
.	.
.	.
c_{mm}^*	$0 \quad 0 \quad \dots 0 \quad 0 \quad \dots 0 \quad \dots r_{m-1,m-1}$

where L is an $m(m-1)/2 \times m(m+1)/2$ matrix which is irrelevant to the value of the determinant of the above matrix, since the corresponding upper right hand partition of the above matrix has all its elements zero.

The Jacobian $J(A^* \rightarrow B, C^*)$ is therefore

$$J(A^* \rightarrow B, C^*) = \prod_{i=1}^m r_{ii}^{m-i}. \quad (2.3.23)$$

Equations (2.3.16), (2.3.20), (2.3.21) and (2.3.23) imply

$$J(X \rightarrow R, \Phi) = \prod_{i=1}^m r_{ii}^{m-i}. \quad (2.3.24)$$

The joint density of R and Φ is then obtained from (2.3.1), (2.3.10) and (2.3.24).

$$h(R, \Phi) = (2\pi)^{-m^2/2} \exp\left\{-\frac{1}{2} \text{tr } RR'\right\} \prod_{i=1}^m r_{ii}^{m-i},$$

on

$$\Phi \in \Omega(m \times m), \quad (2.3.24)$$

$$r_{ii} > 0: \quad i = 1, \dots, m,$$

and

$$-\infty < r_{ij} < +\infty: \quad i = 2, \dots, m; \quad j = 1, \dots, i-1.$$

This density factors into

$$h(R, \Phi) = h_1(\Phi) \left(\prod_{i=1}^m h_i(r_{ii}) \right) \left(\prod_{i=2}^m \prod_{j=1}^{i-1} h_{ij}(r_{ij}) \right), \quad (2.3.25)$$

$$h_1(\Phi) = \prod_{j=1}^m \frac{\Gamma(j/2)}{2\pi^{j/2}} \quad \text{on } \Phi \in \Omega(m \times m), \quad (2.3.26)$$

$$h_i(r_{ii}) = \left[\Gamma\left(\frac{m-i+1}{2}\right) 2^{\frac{m-i-1}{2}} \right]^{-1} r_{ii}^{m-i} \exp\left\{-\frac{1}{2} r_{ii}^2\right\} \\ \text{on } r_{ii} > 0,$$

and

$$h_{ij}(r_{ij}) = (2\pi)^{-1/2} \exp\left\{-\frac{1}{2} r_{ij}^2\right\} \quad \text{on } -\infty < r_{ij} < +\infty.$$

Thus Φ has invariant measure on $\Omega(m \times m)$, and its elements are independent of the elements of R . Furthermore,

$$r_{ii} \wedge \chi_{m-i+1}(0): \quad i = 1, \dots, m,$$

and (2.3.27)

$$r_{ij} \wedge N(0,1): \quad i = 2, \dots, m; \quad j = 1, \dots, i-1.$$

Using (2.3.4) along with the distributional properties of X , R and Φ , the orthospherical integrals can be evaluated. Some examples follow.

$$1. \quad x_{11} = r_{11}\phi_{11}. \quad (2.3.28)$$

Hence,

$$\begin{aligned} E(x_{11}^2) &= E(r_{11}^2 \phi_{11}^2) \\ &= E(r_{11}^2) E(\phi_{11}^2) \\ &= E(r_{11}^2) J_m(2). \end{aligned} \quad (2.3.29)$$

Equations (2.3.1) and (2.3.27) imply that

$$E(x_{11}^2) = 1$$

and

(2.3.30)

$$E(r_{11}^2) = m.$$

Substitution into (2.3.29) yields

$$J_m(2) = \frac{1}{m}. \quad (2.3.31)$$

$$\begin{aligned} 2. \quad E(x_{11}^4) &= E(r_{11}^4 \phi_{11}^4) \\ &= E(r_{11}^4) E(\phi_{11}^4) \\ &= E(r_{11}^4) J_m(4). \end{aligned} \quad (2.3.32)$$

But

$$E(x_{11}^4) = 3$$

and

(2.3.33)

$$E(r_{11}^4) = m(m + 2),$$

implying that

$$J_m(4) = \frac{3}{m(m + 2)}. \quad (2.3.34)$$

$$3. \quad x_{12} = r_{11} \phi_{12}. \quad (2.3.35)$$

From (2.3.28),

$$E(x_{11}^2 x_{12}^2) = E(r_{11}^4 \phi_{11}^2 \phi_{12}^2), \quad (2.3.36)$$

and hence

$$\begin{aligned} E(x_{11}^2) E(x_{12}^2) &= E(r_{11}^4) E(\phi_{11}^2 \phi_{12}^2) \\ &= E(r_{11}^4) J_m(2,2). \end{aligned} \quad (2.3.37)$$

Since x_{11} and x_{12} are independently distributed as $N(0,1)$,

$$J_m(2,2) = \frac{1}{m(m+2)}. \quad (2.3.38)$$

The value of $J_m(2,2)$ could have been obtained from the relationship given by (2.2.22). We have

$$J_m(2) = J_m(4) + (m-1) J_m(2,2). \quad (2.3.39)$$

Substitution of (2.3.31) and (2.3.34) into (2.3.39) yields

$$\frac{1}{m} = \frac{3}{m(m+2)} + (m-1) J_m(2,2), \quad (2.3.40)$$

so that

$$J_m(2,2) = \frac{1}{m(m+2)}. \quad (2.3.41)$$

Other orthospherical integrals may be evaluated in a similar manner. The orthospherical integrals up to order eight, necessary in order to obtain the first four moments of T , are contained in Table 2.1.

TABLE 2.1
Orthospherical Integrals

(The integrals up to order eight not included in this table are equal to zero)

$$\Delta_1 = m(m + 2)$$

$$\Delta_2 = m(m + 2)(m + 4)$$

$$\Delta_3 = m(m + 2)(m + 4)(m + 6)$$

Order 2:

$$J_m(2) = \frac{1}{m}$$

Order 4:

$$J_m(4) = \frac{3}{\Delta_1}$$

$$J_m(2, 2) = \frac{1}{\Delta_1}$$

$$J_m \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \frac{(m+1)}{(m-1)\Delta_1}$$

$$J_m \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = - \frac{1}{(m-1)\Delta_1}$$

Order 6:

$$J_m(6) = \frac{15}{\Delta_2}$$

$$J_m(4, 2) = \frac{3}{\Delta_2}$$

TABLE 2.1—*Continued*

$$J_m \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} = \frac{3(m+3)}{(m-1)\Delta_2}$$

$$J_m \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} = - \frac{3}{(m-1)\Delta_2}$$

$$J_m(2, 2, 2) = \frac{1}{\Delta_2}$$

$$J_m \begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix} = \frac{(m+1)}{(m-1)\Delta_2}$$

$$J_m \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \frac{(m^2 + 3m - 2)}{(m-2)(m-1)\Delta_2}$$

$$J_m \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \frac{1}{(m-1)\Delta_2}$$

$$J_m \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = - \frac{(m+2)}{(m-2)(m-1)\Delta_2}$$

$$J_m \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \frac{2}{(m-2)(m-1)\Delta_2}$$

Order 8:

$$J_m(8) = \frac{105}{\Delta_3}$$

$$J_m(6, 2) = \frac{15}{\Delta_3}$$

$$J_m \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} = \frac{15(m+5)}{(m-1)\Delta_3}$$

TABLE 2.1—*Continued*

$$J_m(4, 4) = \frac{9}{\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = \frac{9(m+3)(m+5)}{(m-1)(m+1)\Delta_3}$$

$$J_m(4, 2, 2) = \frac{3}{\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 2 \\ 2 & 0 \end{pmatrix} = \frac{3(m+1)}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 2 \\ 0 & 2 \end{pmatrix} = \frac{3(m+3)}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \frac{3(m+5)}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 2 \end{pmatrix} = \frac{3(m+3)(m+5)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \frac{3(m^3+8m^2+13m-2)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = -\frac{3}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = -\frac{3(m^2+7m+14)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 3 & 3 \\ 1 & 1 \end{pmatrix} = -\frac{9}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} = -\frac{9(m+3)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \frac{3}{(m-1)\Delta_3}$$

TABLE 2.1—*Continued*

$$J_m \begin{pmatrix} 3 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix} = \frac{3(m+3)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 3 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = - \frac{3(m^2+5m+2)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m(2, 2, 2, 2) = \frac{1}{\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 2 \\ 2 & 0 & 0 \end{pmatrix} = \frac{(m+3)}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \frac{(m+5)}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = \frac{(m^2+4m+15)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 0 \\ 0 & 2 & 2 \end{pmatrix} = \frac{(m+3)^2}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix} = \frac{(m+3)(m+5)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \frac{(m^3+6m^2+3m-6)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix} = \frac{(m+3)(m+5)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \frac{(m^3+8m^2+13m-2)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \frac{(m+3)(m^3+4m^2-11m-2)}{(m-3)(m-2)(m-1)(m+1)\Delta_3}$$

TABLE 2.1—*Continued*

$$J_m \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} = - \frac{(m-3)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = - \frac{1}{(m-1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = - \frac{(m^2+7m+14)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = - \frac{(m^2+3m+6)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 1 & 1 \end{pmatrix} = - \frac{(m+3)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = - \frac{(m^2+5m+2)}{(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} = - \frac{(m+3)}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = - \frac{(m^3+6m^2+3m-6)}{(m-3)(m-2)(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \frac{3}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \frac{1}{(m-1)(m+1)\Delta_3}$$

$$J_m \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \frac{(m^2+6m+16)}{(m-3)(m-2)(m-1)(m+1)\Delta_3}$$

CHAPTER III

MOMENTS OF HOTELLING'S GENERALIZED T_0^2

3.1 Introduction

In Chapter I we defined Hotelling's generalized T_0^2 (see equation (1.1.11)) as

$$\begin{aligned} T_0^2 &= v_2 \sum_{i=1}^{v_1} \underline{Y}_i' S_2^{-1} \underline{Y}_i \\ &= v_2 \operatorname{tr} S_1 S_2^{-1}, \end{aligned} \quad (3.1.1)$$

where, under normal assumptions on the underlying random variables, S_1 and S_2 are $m \times m$ independent random Wishart matrices with

$$S_1 \sim W_m(V, v_1, MM')$$

and

$$(3.1.2)$$

$$S_2 \sim W_m(V, v_2, (0)).$$

It is assumed that V is positive definite. In view of the invariance of the general linear hypothesis and T_0^2 under linear transformations (see Section 1.1), we may simplify the derivation of the moments of T_0^2 .

Since V is positive definite and

$$MM' = \sum_{i=1}^{v_1} \underline{\mu}_i \underline{\mu}_i' \quad (3.1.3)$$

(see equation (1.1.13)) is nonnegative definite, there exists an $m \times m$ matrix, G , satisfying

$$GG' = I$$

and (3.1.4)

$$GMM'G' = \Lambda$$

where Λ is an $m \times m$ diagonal matrix with diagonal elements $\lambda_1, \dots, \lambda_m$ defined in (1.1.23). We may therefore assume, without loss of generality, that in our definition of T_0^2 given by (3.1.1), the column vectors $\underline{y}_i$: $i = 1, \dots, v_1$ and the $m \times m$ matrix S_2 are independently distributed with

$$\underline{y}_i \sim N_m(\underline{\xi}_i, I): \quad i = 1, \dots, v_1, \quad (3.1.5)$$

and

$$S_2 \sim W_m(I, v_2, (0)), \quad (3.1.6)$$

where $\underline{\xi}_i$: $i = 1, \dots, v_1$, satisfy

$$\sum_{i=1}^{v_1} \underline{\xi}_i \underline{\xi}_i' = \Lambda. \quad (3.1.7)$$

Write

$$Y = (\underline{y}_1, \dots, \underline{y}_{v_1}) \quad (3.1.8)$$

and

$$\Xi = (\underline{\xi}_1, \dots, \underline{\xi}_{v_1}) \quad (3.1.9)$$

so that Y is an $m \times v_1$ matrix of random variables and Ξ is an $m \times v_1$ matrix of means, and

$$Y \sim N_{m \times v_1}(\Xi, I \otimes I). \quad (3.1.10)$$

From (3.1.1) we may write

$$\begin{aligned} T_O^2 &= v_2 \sum_{i=1}^{v_1} \Sigma_i' S_2^{-1} \Sigma_i \\ &= v_2 \operatorname{tr} Y' S_2^{-1} Y. \end{aligned} \quad (3.1.11)$$

In this chapter, a method is developed for evaluating the moments of T_O^2 , or, more precisely, the moments of

$$T = \frac{T_O^2}{v_2} \quad (3.1.12)$$

under the assumptions (3.1.6) and (3.1.10) with T_O^2 defined by (3.1.11). The first two moments, $E[T]$ and $E[(T)^2]$, are found explicitly. It is noted that the r th moment of T exists, provided $v_2 > 2r + m - 1$.

3.2 A Method for Evaluating the Moments of T_O^2

We have

$$\begin{aligned} T &= \frac{T_O^2}{v_2} \\ &= \operatorname{tr} Y' S_2^{-1} Y, \end{aligned} \quad (3.2.1)$$

with Y and S_2 distributed as in (3.1.10) and (3.1.6). Let $d_1, \dots, d_m$ be the solutions of the determinantal equation

$$|S_2 - dI| = 0 \quad (3.2.2)$$

so that $d_1, \dots, d_m$ are the latent roots of S_2 . Define D to be an $m \times m$ diagonal matrix with

$$D = \begin{bmatrix} d_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & d_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & d_m \end{bmatrix}. \quad (3.2.3)$$

If $\underline{\ell}_i$ is an m -dimensional random, normed eigenvector associated with d_i , then

$$LS_2L' = D, \quad (3.2.4)$$

where L is the $m \times m$ random matrix satisfying

$$L = (\underline{\ell}_1, \dots, \underline{\ell}_m) \quad (3.2.5)$$

and

$$LL' = I, \quad (3.2.6)$$

so that L is orthogonal. L and D are uniquely defined by (3.2.4), (3.2.5) and (3.2.6) if we require $d_1 > d_2 > \dots > d_m$ and the elements of ℓ_{1j} : $j = 1, \dots, m$, of the first row of L to be positive. The joint density of L and D will be obtained, and it will be shown that the elements of L are independent of the elements of D . The Jacobian, $J(S_2 \rightarrow L, D)$, of the transformation from S_2 to L and D will be obtained by the method used in Section 2.3.

Let δS_2 , δL , and δD be $m \times m$ matrices of elements $(\delta S_2)_{ij}$, $(\delta L)_{ij}$ and $(\delta D)_{ij}$, representing small changes in

the elements of S_2 , L , and D respectively. We require that $(L + \delta L)$ remain orthogonal and $(D + \delta D)$ remain diagonal. Suppose that the changes δL , in L , and δD , in D , cause a change δS_2 in S_2 so that (3.2.4) and (3.2.6) are preserved. Then

$$\begin{aligned} S_2 + \delta S_2 &= (L + \delta L)' (D + \delta D) (L + \delta L) \\ I &= (L + \delta L) (L + \delta L)' . \end{aligned} \quad (3.2.7)$$

By expanding (3.2.7) and dropping terms of order greater than one in the δ 's and using (3.2.4) and (3.2.6), we have

$$\delta S_2 = L'D(\delta L) + L'(\delta D)L + (\delta L)'DL, \quad (3.2.8)$$

and

$$0 = L(\delta L)' + (\delta L)L'.$$

As in Section 2.3

$$J(S_2 \rightarrow L, D) = J(\delta S_2 \rightarrow \delta L, \delta D) \quad (3.2.9)$$

where $J(\delta S_2 \rightarrow \delta L, \delta D)$ is the Jacobian of the transformation from δS_2 to δL and δD defined by (3.2.8), in which L and D are now assumed to be fixed $m \times m$ matrices. For convenience, denote δS_2 by A , δL by B , and δD by C . Then from (3.2.8) and (3.2.9) we seek the Jacobian $J(A \rightarrow B, C)$ defined by

$$\begin{aligned} A &= L'DB + L'CL + B'DL \\ \text{and} \end{aligned} \quad (3.2.10)$$

$$0 = LB' + BL'.$$

Define $m \times m$ matrices A^* and B^* by

$$A^* = LAL'$$

and

$$(3.2.11)$$

$$B^* = BL'.$$

Using (3.2.11), equations (3.2.10) may be written as

$$A^* = DB^* + C + B^{*'}D$$

and

$$(3.2.12)$$

$$0 = B^{*'} + B^*,$$

so that B^* is an asymmetric matrix. Then

$$J(A \rightarrow B, C) = J(A \rightarrow A^*) J(A^* \rightarrow B^*, C) J(B^* \rightarrow B) \quad (3.2.13)$$

where $J(A \rightarrow A^*)$ and $J(B^* \rightarrow B)$ are the Jacobians of the transformations defined by equations (3.2.11), and $J(A^* \rightarrow B^*, C)$ is the Jacobian of the transformation defined by equations (3.2.12). Equations (3.2.11) represent orthogonal transformations and hence

$$J(A \rightarrow A^*) = 1$$

and

$$(3.2.14)$$

$$J(B^* \rightarrow B) = 1.$$

Equation (3.2.13) then reduces to

$$J(A \rightarrow B, C) = J(A^* \rightarrow B^*, C). \quad (3.2.15)$$

Direct evaluation of $J(A^* \rightarrow B^*, C)$ is straightforward.

Recall that $C = \beta D$ and hence is diagonal, B^* is asymmetric and A^* is symmetric.

$$J(A^* \rightarrow B^*, C) = \left\| \frac{\partial (a_{11}^* a_{22}^* \cdots a_{mm}^* a_{21}^* a_{31}^* a_{32}^* \cdots a_{m1}^* a_{m2}^*)}{\partial (c_{11} \ c_{22} \ \cdots c_{mm} \ b_{21}^* b_{31}^* b_{32}^* \cdots b_{m1}^* b_{m2}^*)} \right\| . \quad (3.2.16)$$

Laying out the matrix of derivatives in tabular form, we have

$$\begin{bmatrix}
 c_{11} & a_{11}^* & a_{22}^* & \dots & a_{mm}^* & a_{21}^* & a_{31}^* & a_{32}^* & \dots & a_{m1}^* & a_{m2}^* & \dots & a_{m,m-1}^* \\
 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
 c_{22} & 0 & 1 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 c_{mm} & 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
 b_{21}^* & 0 & 0 & \dots & 0 & d_2-d_1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\
 b_{31}^* & 0 & 0 & \dots & 0 & 0 & d_3-d_1 & 0 & \dots & 0 & 0 & \dots & 0 \\
 b_{32}^* & 0 & 0 & \dots & 0 & 0 & 0 & d_3-d_2 & \dots & 0 & 0 & \dots & 0 \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 b_{m1}^* & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & d_m-d_1 & 0 & \dots & 0 \\
 b_{m2}^* & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & d_m-d_2 & \dots & 0 \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 . & . & . & . & . & . & . & . & . & . & . & . & . \\
 b_{m,m-1}^* & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & d_m-d_{m-1}
 \end{bmatrix}$$

The Jacobian $J(A^* \rightarrow B^*, C)$ is the absolute value of the determinant of the above matrix and is

$$J(A^* \rightarrow B^*, C) = \prod_{1 \leq i < j \leq m} (d_i - d_j). \quad (3.2.17)$$

Thus, from equations (3.2.9), (3.2.15) and (3.2.16),

$$J(S_2 \rightarrow L, D) = \prod_{1 \leq i < j \leq m} (d_i - d_j). \quad (3.2.18)$$

From (3.1.6), the density on S_2 is

$$\begin{aligned} f(S_2) &= \left[2^{mv_2/2} \pi^{m(m-1)/4} \prod_{j=1}^m \Gamma\left(\frac{v_2 + 1 - j}{2}\right) \right]^{-1} \\ &\times |S_2|^{(v_2 - m - 1)/2} \exp\left[-\frac{1}{2} \text{tr } S_2\right] \end{aligned} \quad (3.2.19)$$

on S_2 positive definite. The joint density of L and D , $g(L, D)$, is, from equation (3.2.4)

$$g(L, D) = f(L'DL) J(S_2 \rightarrow L, D) \quad (3.2.20)$$

where $f(\cdot)$ is the density of (3.2.19). Recalling that L is orthogonal, we find that

$$\begin{aligned} g(L, D) &= \left[2^{mv_2/2} \pi^{m(m-1)/4} \prod_{j=1}^m \Gamma\left(\frac{v_2 + 1 - j}{2}\right) \right]^{-1} \\ &\times \left[\prod_{j=1}^m d_j^{(v_2 - m - 1)/2} \exp\left(-\frac{1}{2} d_j\right) \right] \prod_{1 \leq i < j \leq m} (d_i - d_j) \end{aligned} \quad (3.2.21)$$

on L satisfying $LL' = I$, $\ell_{1j} > 0$: $j = 1, \dots, m$ and $d_1 > d_2 > \dots > d_m > 0$. The density of (3.2.21) factors into

$$g(L, D) = g_1(L) g_2(D) \quad (3.2.22)$$

where

$$\begin{aligned} g_1(L) &= \prod_{j=1}^m \Gamma\left(\frac{j}{2}\right) / \pi^{j/2} \\ &\quad \text{if } LL' = I, \ell_{1j} > 0: j = 1, \dots, m \\ &= 0 \quad \text{otherwise,} \end{aligned} \quad (3.2.23)$$

and

$$\begin{aligned} g_2(D) &= \left[2^{mv_2/2} \prod_{j=1}^m \Gamma\left(\frac{v_2 + 1}{2} - \frac{j}{2}\right) \Gamma\left(\frac{j}{2}\right) \right]^{-1} \pi^{m/2} \\ &\quad \times \left[\prod_{j=1}^m d_j^{(v_2 - m - 1)/2} \exp\left(-\frac{1}{2} d_j\right) \right] \prod_{1 \leq i < j \leq m} (d_i - d_j) \\ &\quad \text{if } d_1 > d_2 > \dots > d_m > 0 \\ &= 0 \quad \text{otherwise.} \end{aligned} \quad (3.2.24)$$

Hence the elements of L are independent of the elements of D .

Suppose that K is any $m \times m$ random orthogonal matrix satisfying

$$KS_2K' = D. \quad (3.2.25)$$

Write

$$K = (\underline{k}_1, \underline{k}_2, \dots, \underline{k}_m) \quad (3.2.26)$$

where $\underline{k}_i$ is an m -dimensional, random, normed eigenvector associated with d_i . Comparing (3.2.26) with (3.2.25), we have

$$P(\underline{k}_i = \underline{\ell}_i) = \frac{1}{2} : i = 1, \dots, m$$

and (3.2.27)

$$P(\underline{k}_i = -\underline{\ell}_i) = \frac{1}{2} : i = 1, \dots, m.$$

The measure on K is obtained from (3.2.23) and (3.2.27).

It is in fact

$$\begin{aligned} h(K) &= \frac{1}{2^m} \prod_{j=1}^m \Gamma\left(\frac{j}{2}\right) / \pi^{j/2} && \text{if } KK' = I \\ &= 0 && \text{otherwise} \end{aligned} \quad (3.2.28)$$

so that K has invariant measure on $\Omega(m \times m)$ and is independent of D .

With Y defined by (3.1.10) and K defined by (3.2.28), let Z be an $m \times v_1$ matrix of random variables satisfying

$$Z = KY. \quad (3.2.29)$$

From (3.1.10), the distribution of Z , conditional on K , is

$$Z|K \sim N_{m \times v_1}(KE, I \otimes I). \quad (3.2.30)$$

The unconditional density on Z can be obtained from (3.2.30) and is

$$f(Z) = \int_{K \in \Omega(m \times m)} (2\pi)^{-mv_1/2} \exp \left[-\frac{1}{2} \text{tr}(Z - KE)(Z - KE)' \right] \\ \times \prod_{j=1}^m \frac{\Gamma(\frac{j}{2})}{2\pi^{j/2}} dK. \quad (3.2.31)$$

It should be noted that since Y and S_2 are independent, Y , K and D are mutually independent. This implies the independence of Z and D .

Using equations (3.2.1), (3.2.25) and (3.2.29), we have

$$T = \text{tr } Y' S_2^{-1} Y \\ = \text{tr}(Z' K)(K' D^{-1} K)(K' Z) \\ = \text{tr } Z' D^{-1} Z \\ = \sum_{i=1}^m \sum_{j=1}^{v_1} \frac{z_{ij}^2}{d_i}. \quad (3.2.32)$$

Writing

$$u_i^2 = \sum_{j=1}^{v_1} z_{ij}^2, \quad (3.2.33)$$

then

$$T = \sum_{i=1}^m \frac{u_i^2}{d_i}. \quad (3.2.34)$$

This representation of T , along with the results of Chapter II, and the distributional properties of Z and D , can be used to evaluate the moments of T .

3.3 The First Two Moments of T

Using the representation of T given in equation (3.2.39) the first two moments of T are

$$\begin{aligned} E(T) &= E \left[\sum_{i=1}^m \frac{u_i^2}{d_i} \right] \\ &= \sum_{i=1}^m E(u_i^2) E\left(\frac{1}{d_i}\right), \end{aligned} \quad (3.3.1)$$

and

$$\begin{aligned} E[(T)^2] &= E \left[\sum_{i=1}^m \frac{u_i^2}{d_i} \right]^2 \\ &= E \left[\sum_{i=1}^m \frac{u_i^4}{d_i^2} + \sum_{i \neq j} \frac{u_i^2 u_j^2}{d_i d_j} \right] \\ &= \sum_{i=1}^m E(u_i^4) E\left(\frac{1}{d_i^2}\right) + \sum_{i \neq j} E(u_i^2 u_j^2) E\left(\frac{1}{d_i d_j}\right). \end{aligned} \quad (3.3.2)$$

From equations (3.2.30) and (3.2.33), u_i^2 conditional on K is the sum of v_1 independent normal random variables.

Thus

$$u_i^2 \Big| K \sim \chi_{v_1}^2 \left[(KAK')_{ii} \right]: \quad i = 1, \dots, m \quad (3.3.3)$$

where from (3.1.7)

$$\Lambda = \Xi \Xi'. \quad (3.3.4)$$

For convenience, let

$$\begin{aligned} \gamma_i &= (\Lambda \Lambda')_{ii} \\ &= \sum_{j=1}^m k_{ij}^2 \lambda_i \end{aligned} \quad (3.3.5)$$

since Λ is diagonal. Then

$$u_i^2 \Big| K \sim \chi_{v_1}^2(\gamma_i): \quad i = 1, 2, \dots, m. \quad (3.3.6)$$

Furthermore, conditional on K , u_i^2 and u_j^2 are independent. Hence, using (3.3.6)

$$\begin{aligned} E(u_i^2) &= E_K E_{u_i^2 | K}(u_i^2) \\ &= E_K(v_1 + \gamma_i), \end{aligned} \quad (3.3.7)$$

$$\begin{aligned} E(u_i^4) &= E_K E_{u_i^2 | K}(u_i^4) \\ &= E_K[v_1(v_1 + 2) + 2\gamma_i(v_1 + 2) + \gamma_i^2], \end{aligned} \quad (3.3.8)$$

and

$$\begin{aligned} E(u_i^2 u_j^2) &= E_K E_{u_i, u_j | K}(u_i^2 u_j^2) \\ &= E_K E_{u_i | K}(u_i^2) E_{u_j | K}(u_j^2) \\ &= E_K[v_1^2 + v_1(\gamma_i + \gamma_j) + \gamma_i \gamma_j]. \end{aligned} \quad (3.3.9)$$

Recalling that K has invariant measure on $\Omega(m \times m)$, we may use the results of Chapter II to evaluate the moments of γ_i and the product moment, $E_K(\gamma_i \gamma_j)$ where γ_i is defined in (3.3.5).

$$\begin{aligned} E_K(\gamma_i) &= E_K \left(\sum_{j=1}^m k_{ij}^2 \lambda_j \right) \\ &= J_m(2) \sum_{j=1}^m \lambda_j \\ &= \frac{1}{m} \sum_{j=1}^m \lambda_j, \end{aligned} \quad (3.3.10)$$

$$\begin{aligned} E_K(\gamma_i^2) &= E_K \left(\sum_{j=1}^m k_{ij}^2 \lambda_j \right)^2 \\ &= E_K \left(\sum_{j=1}^m k_{ij}^4 \lambda_j^2 + \sum_{\ell \neq j} k_{i\ell}^2 k_{ij}^2 \lambda_\ell \lambda_j \right) \\ &= J_m(4) \sum_{j=1}^m \lambda_j^2 + J_m(2,2) \sum_{\ell \neq j} \lambda_\ell \lambda_j \\ &= \left(\frac{3}{m(m+2)} \right) \sum_{j=1}^m \lambda_j^2 + \left(\frac{1}{m(m+2)} \right) \sum_{\ell \neq j} \lambda_\ell \lambda_j, \end{aligned} \quad (3.3.11)$$

and

$$\begin{aligned} E_K(\gamma_i \gamma_j) &= E_K \left(\sum_{\ell=1}^m k_{i\ell}^2 \lambda_\ell \right) \left(\sum_{\ell=1}^m k_{j\ell}^2 \lambda_\ell \right) \\ &= E_K \left(\sum_{\ell=1}^m k_{i\ell}^2 k_{j\ell}^2 \lambda_\ell^2 + \sum_{\ell \neq \ell'} k_{i\ell}^2 k_{j\ell'}^2 \lambda_\ell \lambda_{\ell'} \right) \end{aligned}$$

$$\begin{aligned}
 &= J_m(2, 2) \sum_{\ell=1}^m \lambda_{\ell}^2 + J_m \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \sum_{\ell \neq \ell} \lambda_{\ell} \lambda_{\ell}, \\
 &= \frac{1}{m(m+2)} \sum_{\ell=1}^m \lambda_{\ell}^2 + \frac{(m+1)}{(m-1)m(m+2)} \sum_{\ell \neq \ell} \lambda_{\ell} \lambda_{\ell},
 \end{aligned}
 \tag{3.3.12}$$

Substitution of equations (3.3.10), (3.3.11) and (3.3.12) into (3.3.7), (3.3.8) and (3.3.9) yields

$$E(u_i^2) = v_1 + \frac{1}{m} \sum_{j=1}^m \lambda_j^2, \tag{3.3.13}$$

$$\begin{aligned}
 E(u_i^4) &= v_1(v_1 + 2) + (2v_1 + 4) \frac{1}{m} \sum_{j=1}^m \lambda_j \\
 &\quad + \frac{3}{m(m+2)} \sum_{j=1}^m \lambda_j^2 + \frac{1}{m(m+2)} \sum_{i \neq j} \lambda_i \lambda_j,
 \end{aligned}
 \tag{3.3.14}$$

and

$$\begin{aligned}
 E(u_i^2 u_j^2) &= v_1^2 + \frac{2v_1}{m} \sum_{j=1}^m \lambda_j^2 \\
 &\quad + \frac{m+1}{(m-1)m(m+2)} \sum_{i \neq j} \lambda_i \lambda_j.
 \end{aligned}
 \tag{3.3.15}$$

Upon substitution of (3.3.13), (3.3.14) and (3.3.15) into equations (3.3.1) and (3.3.2), the first two moments of T are

$$E(T) = \left(v_1 + \frac{1}{m} \sum_{j=1}^m \lambda_j^2 \right) E \sum_{i=1}^m \left(\frac{1}{d_i} \right) \tag{3.3.16}$$

and

$$\begin{aligned}
 E((T)^2) &= \left[v_1(v_1 + 2) + \frac{2v_1 + 4}{m} \sum_{j=1}^m \lambda_j \right. \\
 &\quad \left. + \frac{5}{m(m+2)} \sum_{j=1}^m \lambda_j^2 + \frac{1}{m(m+2)} \sum_{i \neq j} \lambda_i \lambda_j \right] \\
 &\times E \left[\sum_{i=1}^m \frac{1}{d_i^2} \right] + \left[v_1^2 + \frac{2v_1}{m} \sum_{j=1}^m \lambda_j \right. \\
 &\quad \left. + \frac{1}{m(m+2)} \sum_{j=1}^m \lambda_j^2 + \frac{m+1}{(m-1)m(m+2)} \sum_{i \neq j} \lambda_i \lambda_j \right] \\
 &\times E \left[\sum_{i \neq j} \frac{1}{d_i d_j} \right]. \tag{3.3.17}
 \end{aligned}$$

Evaluation of the terms

$$E \left[\sum_{i=1}^m \frac{1}{d_i} \right], \quad E \left[\sum_{i=1}^m \frac{1}{d_i^2} \right]$$

and

$$E \left[\sum_{i \neq j} \frac{1}{d_i d_j} \right]$$

involve some special problems. These will be dealt with separately in the next section. We list these expectations here, however, in order to obtain the moments of T:

$$E \left[\sum_{i=1}^m \frac{1}{d_i} \right] = \frac{m}{v_2 - m - 1}, \tag{3.3.18}$$

$$E \left(\sum_{i=1}^m \frac{1}{d_i^2} \right) = \frac{m(v_2-1)}{(v_2-m)(v_2-m-1)(v_2-m-3)}, \quad (3.3.19)$$

and

$$E \left(\sum_{i \neq j} \frac{1}{d_i d_j} \right) = \frac{m(m-1)}{(v_2-m)(v_2-m-1)}. \quad (3.3.20)$$

Substituting these expectations into equations (3.3.16) and (3.3.17), we obtain

$$E(T) = \frac{mv_1 + \sum_{i=1}^m \lambda_i}{v_2-m-1} \quad (3.3.21)$$

and

$$\begin{aligned} & (v_2-m)(v_2-m-1)(v_2-m-3) E\{(T)^2\} \\ &= mv_1^2(mv_2-m^2-2m+2) + 2mv_1(v_2-1) \\ &+ \{2v_1(mv_2-m^2-2m+2) + 4(v_2-1)\} \sum_{i=1}^m \lambda_i \\ &+ (v_2-m-2) \left(\sum_{i=1}^m \lambda_i \right)^2 + 2 \sum_{i=1}^m \lambda_i^2. \end{aligned} \quad (3.3.22)$$

Numerous checks are available on the moments of T . The distribution of T , and hence its moments, is known in the following special cases. If $m = 1$,

$$T \sim \frac{\chi_{v_1}^2(\lambda)}{\chi_{v_2}^2(0)}. \quad (3.3.23)$$

Putting $m = 1$ in equations (3.3.21) and (3.3.22), the first two moments of T reduce to the first two moments of a random variable distributed as in (3.3.23). In the case of $v_1 = 1$, T_0^2 reduces to Hotelling's T^2 defined in (1.2.1). Hence when $v_1 = 1$,

$$T = \frac{T^2}{v^2}. \quad (3.3.24)$$

The distribution of Hotelling's T^2 is known to be [14]

$$T^2 \sim \frac{v_2 \chi_m^2(\lambda)}{\chi_{v_2 - m + 1}^2(0)}. \quad (3.3.25)$$

It should be noted that when $v_1 = 1$, there is at most one nonzero solution of

$$|\Lambda - \lambda I| = 0. \quad (3.3.26)$$

Then when $v_1 = 1$

$$T \sim \frac{\chi_m^2(\lambda)}{\chi_{v_2 - m + 1}^2(0)}. \quad (3.3.27)$$

Putting $v_1 = 1$ in equations (3.3.21) and (3.3.22), we see that the first two moments of T reduce to those of a random variable distributed as in (3.3.27). The limiting distribution of T_0^2 , as v_2 tends to infinity, is known to be (see [15] or [26])

$$\lim_{v_2 \rightarrow \infty} T_0^2 \sim \chi_{mv_1}^2 \left(\sum_{i=1}^m \lambda_i \right). \quad (3.3.28)$$

In equation (3.3.21), if we multiply by v_2 and take the limit we obtain

$$E \left[\lim_{v_2 \rightarrow \infty} T_0^2 \right] = m v_1 + \sum_{i=1}^m \lambda_i. \quad (3.3.29)$$

Similarly, multiplying (3.3.22) by v_2^2 and taking the limit, we have

$$E \left[\lim_{v_2 \rightarrow \infty} T_0^4 \right] = m^2 v_1^2 + 2m v_1 + 2(m v_1 + 2) \sum_{i=1}^m \lambda_i + \left(\sum_{i=1}^m \lambda_i \right)^2. \quad (3.3.30)$$

Equations (3.3.29) and (3.3.30) are the first two moments of a random variable distributed as in (3.3.28). As a final check on the moments of T , it is known that the distribution of T is invariant under the transformation (see [1])

$$(m, v_1, v_2) \rightarrow (v_1, m, v_1 + v_2 - m). \quad (3.3.31)$$

Simple substitution shows that the moments of T given in (3.3.21) and (3.3.22) are in fact invariant under this transformation.

In the null case when $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$, the moments of T reduce to

$$E(T | \lambda_1 = \dots = \lambda_m = 0) = \frac{m v_1}{v_2 - m - 1}, \quad (3.3.32)$$

and

$$\begin{aligned} & (v_2 - m)(v_2 - m - 1)(v_2 - m - 3) E\left((T)^2 \mid \lambda_1 = \cdots = \lambda_m = 0\right) \\ & = mv_1^2(mv_2 - m^2 - 2m + 2) + 2mv_1(v_2 - 1). \end{aligned} \quad (3.3.33)$$

3.4 Evaluation of Certain Expectations Used in Obtaining the Moments of T

In Section 3.3, we listed the values of $E\left(\sum_{i=1}^m \frac{1}{d_i}\right)$

$$E\left(\sum_{i=1}^m \frac{1}{d_i^2}\right) \text{ and } E\left(\sum_{i \neq j} \frac{1}{d_i d_j}\right)$$

in order to evaluate the moments of T (see equations (3.3.18), (3.3.19) and (3.3.20)). We derive these expectations here. Recall that $d_1, \dots, d_m$ are the latent roots of S_2 . Thus

$$\sum_{i=1}^m \frac{1}{d_i} = \text{tr } S_2^{-1}, \quad (3.4.1)$$

$$\sum_{i=1}^m \frac{1}{d_i^2} = \text{tr } (S_2^{-1})^2, \quad (3.4.2)$$

and

$$\sum_{i \neq j} \frac{1}{d_i d_j} = (\text{tr } S_2^{-1})^2 - \text{tr } (S_2^{-1})^2. \quad (3.4.3)$$

We shall need some results in matrix algebra and some well known properties of the Wishart random matrix (see [1] or [32]). Let A be an $m \times m$ nonsingular matrix and partition A into

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array} \right] \quad (3.4.4)$$

where A_{ij} is a $p_i \times p_j$ matrix with $p_1 + p_2 = m$. Then

$$|A| = |A_{11}| |A_{22} - A_{21}A_{11}^{-1}A_{12}|,$$

or (3.4.5)

$$= |A_{22}| |A_{11} - A_{12}A_{22}^{-1}A_{21}|.$$

Suppose that B is an $m \times m$ matrix satisfying

$$B = A^{-1}. \quad (3.4.6)$$

Partition B in the same manner as A ,

$$B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array} \right]. \quad (3.4.7)$$

Then

$$B_{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}$$

and (3.4.8)

$$B_{22} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}.$$

With A defined and partitioned as above, suppose that

$$A \sim W_m(I, \nu, (0)). \quad (3.4.9)$$

Then A_{11} and A_{22} are independently distributed and

$$A_{ii} \sim W_{p_i}(I, v, (0)): \quad i = 1, 2. \quad (3.4.10)$$

Also, A_{ii} is distributed independently of $(A_{jj} - A_{ji}A_{ii}^{-1}A_{ij})$, with

$$(A_{jj} - A_{ji}A_{ii}^{-1}A_{ij}) \sim W_{p_j}(I, v - p_i, (0)). \quad (3.4.11)$$

Furthermore, $(A_{jj} - A_{ji}A_{ii}^{-1}A_{ij})$ is independent of $A_{ji}A_{ii}^{-1}A_{ij}$ and

$$A_{ji}A_{ii}^{-1}A_{ij} \sim W_{p_j}(I, p_i, (0)). \quad (3.4.12)$$

It is important to note that the Wishart distribution with $m = 1$ on v degrees of freedom is identical to the χ^2 -distribution on v degrees of freedom. Thus if

$$u \sim W_1(1, v, 0),$$

then

$$(3.4.13)$$

$$u \sim \chi_v^2(0).$$

With these results, we can proceed to evaluate $E(\text{tr } S_2^{-1})$, $E\{\text{tr}(S_2^{-1})^2\}$ and $E\{(\text{tr } S_2^{-1})^2 - \text{tr}(S_2^{-1})^2\}$. For convenience, we shall work with the corresponding expectations in terms of A , where A is defined by (3.4.9). We require $E(\text{tr } A^{-1})$, $E\{\text{tr}(A^{-1})^2\}$ and $E\{(\text{tr } A^{-1})^2 - \text{tr}(A^{-1})^2\}$. Let

$$B = A^{-1}. \quad (3.4.14)$$

Then, by symmetry,

$$\begin{aligned}
 E(\text{tr } A^{-1}) &= E(\text{tr } B) \\
 &= E\left[\sum_{i=1}^m b_{ii}\right] \\
 &= mE(b_{11}).
 \end{aligned} \tag{3.4.15}$$

From equation (3.4.8) with $p_1 = 1$ and $p_2 = m - 1$,

$$b_{11} = (a_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}. \tag{3.4.16}$$

Thus using (3.4.11) and (3.4.13)

$$b_{11} \sim \frac{1}{\chi_{v^2-m+1}(0)}, \tag{3.4.17}$$

and hence

$$E(\text{tr } A^{-1}) = \frac{m}{(v-m-1)}. \tag{3.4.18}$$

The evaluation of $E\{(\text{tr } \lambda^{-1})^2 - \text{tr}(A^{-1})^2\}$ is similar. By symmetry

$$\begin{aligned}
 E\{(\text{tr } A^{-1})^2 - \text{tr}(A^{-1})^2\} &= E\{(\text{tr } B)^2 - \text{tr}(B)^2\} \\
 &= E\left[\left(\sum_{i=1}^m b_{ii}\right)^2 - \sum_{i=1}^m \sum_{j=1}^m b_{ij}^2\right] \\
 &= E\left[\sum_{i \neq j} (b_{ii}b_{jj} - b_{ij}^2)\right] \\
 &= m(m-1) E(b_{11}b_{22} - b_{12}^2).
 \end{aligned} \tag{3.4.19}$$

With $p_1 = 2$ and $p_2 = m - 2$ and using (3.4.8), we have

$$\begin{aligned} b_{11}b_{22} - b_{12}^2 &= |B_{11}| \\ &= |(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}|. \end{aligned} \quad (3.4.20)$$

Therefore,

$$E((\text{tr } A^{-1})^2 - \text{tr}(A^{-1})^2) = m(m-1) E \left[\frac{1}{|A_{11} - A_{12}A_{22}^{-1}A_{21}|} \right]. \quad (3.4.21)$$

Denote $(A_{11} - A_{12}A_{22}^{-1}A_{21})$ by

$$\begin{aligned} (A_{11} - A_{12}A_{22}^{-1}A_{21}) &= C \\ &= \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \end{aligned} \quad (3.4.22)$$

so that from (3.4.11)

$$C \sim W_2(I, v-m+2, (0)). \quad (3.4.23)$$

Now from equations (3.4.5), (3.4.10) and (3.4.11),

$$|C| = c_{11} \left[c_{22} - \frac{c_{12}^2}{c_{11}} \right] \quad (3.4.24)$$

with

$$c_{11} \sim \chi_{v-m+2}^2(0)$$

and

(3.4.25)

$$\left(c_{22} - \frac{c_{12}^2}{c_{11}} \right) \sim \chi_{v-m+1}^2(0).$$

This implies that

$$\begin{aligned} E((\text{tr } A^{-1})^2 - \text{tr}(A^{-1})^2) &= m(m-1) E\left(\frac{1}{\chi_{v-m+2}^2}\right) E\left(\frac{1}{\chi_{v-m+1}^2}\right) \\ &= \frac{m(m-1)}{(v-m)(v-m-1)}. \end{aligned} \quad (3.4.26)$$

The evaluation of $E(\text{tr}(A^{-1})^2)$ can be carried out using (3.4.26) and evaluating $E(\text{tr } A^{-1})^2$. Again by symmetry we have

$$\begin{aligned} E(\text{tr } A^{-1})^2 &= E(\text{tr } B)^2 \\ &= E\left(\sum_{i=1}^m b_{ii}\right)^2 \\ &= E\left(\sum_{i=1}^m b_{ii}^2 + \sum_{i \neq j} b_{ii} b_{jj}\right) \\ &= mE(b_{11}^2) + m(m-1) E(b_{11} b_{22}). \end{aligned} \quad (3.4.27)$$

From (3.4.17),

$$E(b_{11}^2) = \frac{1}{(v-m-1)(v-m-5)}. \quad (3.4.28)$$

To evaluate $E(b_{11} b_{22})$ consider (3.4.7) and (3.4.8) with $p_1 = 2$ and $p_2 = m - 2$. We have

$$\begin{aligned}
 B_{11} &= \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \\
 &= (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} \\
 &= C^{-1}
 \end{aligned} \tag{3.4.29}$$

where C is defined in (3.4.22) and distributed as in (3.4.23). Another application of (3.4.8) yields

$$b_{11} = \left(c_{11} - \frac{c_{12}^2}{c_{22}} \right)^{-1}$$

and (3.4.30)

$$b_{22} = \left(c_{22} - \frac{c_{12}^2}{c_{11}} \right)^{-1}$$

Thus

$$b_{11}b_{22} = \left[c_{11}c_{22} \left(1 - \frac{c_{12}^2}{c_{11}c_{22}} \right)^2 \right]^{-1}. \tag{3.4.31}$$

The remarks preceding (3.4.10) along with (3.4.23) imply the independence of c_{11} and c_{22} . Also, the remarks preceding (3.4.11) imply the independence of c_{11} and $c_{22} - \frac{c_{12}^2}{c_{11}}$. Hence c_{11} is independent of $\left(1 - \frac{c_{12}^2}{c_{11}c_{22}} \right)$. Similarly, c_{22} is independent of $\left(1 - \frac{c_{12}^2}{c_{11}c_{22}} \right)$ and, hence, so is the product $c_{11}c_{22}$. This implies that

$$\begin{aligned}
 E(b_{11}b_{22}) &= E\left(\frac{1}{c_{11}}\right) E\left(\frac{1}{c_{22}}\right) E\left(1 - \frac{c_{12}^2}{c_{11}c_{22}}\right)^{-2} \\
 &= E\left(\frac{1}{c_{11}}\right) E\left(\frac{1}{c_{22}}\right) E\left(\frac{c_{11}}{c_{11} - \frac{c_{12}^2}{c_{22}}}\right)^2. \quad (3.4.32)
 \end{aligned}$$

The expectations $E\left(\frac{1}{c_{11}}\right)$ and $E\left(\frac{1}{c_{22}}\right)$ can be evaluated using (3.4.10), (3.4.13) and (3.4.23), yielding

$$E\left(\frac{1}{c_{11}}\right) = \frac{1}{v-m}$$

and

(3.4.33)

$$E\left(\frac{1}{c_{22}}\right) = \frac{1}{v-m}.$$

To evaluate

$$E\left(\frac{c_{11}}{c_{11} - \frac{c_{12}^2}{c_{22}}}\right)^2,$$

we need to make use of (3.4.12). Let

$$u_1 = c_{11} - \frac{c_{12}^2}{c_{22}}$$

and

(3.4.34)

$$u_2 = \frac{c_{12}^2}{c_{22}}.$$

From (3.4.12) and (3.4.23), u_1 and u_2 are independently distributed as

$$u_1 \sim \chi_{v-m+1}^2(0)$$

and (3.4.35)

$$u_2 \sim \chi_1^2(0).$$

Therefore

$$\begin{aligned} \frac{c_{11}}{c_{11} - \frac{c_{12}^2}{22}} &= \frac{u_1 + u_2}{u_1} \\ &\sim \frac{1}{\beta e(1, v-m+1)}. \end{aligned} \quad (3.4.36)$$

Equation (3.4.36) yields

$$E \left[\frac{c_{11}}{c_{11} - \frac{c_{12}^2}{22}} \right]^2 = \frac{(v-m)(v-m-2)}{(v-m-1)(v-m-3)}. \quad (3.4.37)$$

Substitution of (3.4.33) and (3.4.37) into (3.4.32) gives

$$E(b_{11}b_{22}) = \frac{(v-m-2)}{(v-m)(v-m-1)(v-m-3)}, \quad (3.4.38)$$

and we obtain from (3.4.27), (3.4.28) and (3.4.38)

$$E(\text{tr } A^{-1})^2 = \frac{m(mv-m^2-2m+2)}{(v-m)(v-m-1)(v-m-3)}. \quad (3.4.39)$$

Finally substituting (3.4.39) into (3.4.26) and solving for $E(\text{tr}(A^{-1})^2)$, we find that

$$E(\text{tr}(A^{-1})^2) = \frac{m(v-1)}{(v-m)(v-m-1)(v-m-3)}. \quad (3.4.40)$$

Putting $A = S_2$ and replacing v by v_2 in equations (3.4.18), (3.4.26) and (3.4.40) we obtain the equations

$$E(\text{tr } S_2^{-1}) = \frac{m}{v_2 - m - 1} \quad (3.4.41)$$

$$E(\text{tr}(S_2^{-1})^2) = \frac{m(v_2 - 1)}{(v_2 - m)(v_2 - m - 1)(v_2 - m - 3)} \quad (3.4.42)$$

and

$$E((\text{tr } S_2^{-1})^2 - \text{tr}(S_2^{-1})^2) = \frac{m(m-1)}{(v_2 - m)(v_2 - m - 1)} \quad (3.4.43)$$

which are the same as equations (3.3.18), (3.3.19) and (3.3.20).

3.5 Summary

The representation of T given in (3.2.34),

$$T = \sum_{i=1}^m \frac{u_i^2}{d_i},$$

can be used to evaluate higher order moments of T . The expectations involving the elements $u_1^2, \dots, u_m^2$, can be evaluated using (3.3.6) and the orthospherical integrals. The orthospherical integrals necessary for the first four moments of T are tabulated at the end of Chapter II. The expectations involving elements of D can be found using simple properties of the Wishart distribution similar to those in Section 3.4.

It should be noted that the second moment of T depends upon

$$\sum_{i=1}^m \lambda_i^2.$$

This refutes the claim by Mikhail [27] and Srivastava [33] that the power of T depends upon $\lambda_1, \dots, \lambda_m$ only through

$$\sum_{i=1}^m \lambda_i.$$

CHAPTER IV

APPROXIMATIONS TO THE DISTRIBUTION OF
HOTELLING'S GENERALIZED T_0^2

4.1 Introduction

It was pointed out in Chapter III that the distribution of Hotelling's generalized T_0^2 is well known in certain special instances. These are the cases when $m = 1$ for all v_1 and v_2 ; when $v_1 = 1$ for all m and $v_2 \geq m$; and when v_2 goes to infinity for all m and v_1 . In addition to these special cases, Hotelling [13] derived the null distribution of T_0^2 for $m = 2$, all v_1 and $v_2 \geq 2$. With the exceptions of these few instances, the distribution of Hotelling's generalized T_0^2 is customarily in series form in terms of the zonal polynomials as described by James [21] and Constantine [3]. Constantine [4] gives a series representation of the distribution of T . His series is convergent, however, only when $|T| < 1$. In 1968, Davis [6] showed that the null distribution of T satisfies a linear homogeneous differential equation of order m , and that Constantine's series reduces to the solution of this equation in the case $|T| < 1$. Davis [7] uses this result to obtain the upper 5% and 1% points of T_0^2/v_1 for $m = 3, 4$

and various values of v_1 and v_2 which he claims are generally accurate to five significant figures.

In 1954, Pillai [29] suggested that the distribution of T be approximated by

$$T \hat{=} \frac{m^2 v_1}{m(v_2 - m - 1) + 2} F_{mv_1, m(v_2 - m - 1) + 2}. \quad (4.1.1)$$

It should be pointed out that this approximation becomes exact for $m = 1$ as well as for the limiting case when v_2 tends to infinity. However, the approximation is not exact in the case $v_1 = 1$, nor is it invariant under the transformation

$$(m, v_1, v_2) \rightarrow (v_1, m, v_1 + v_2 - m).$$

In this chapter three approximations to the distribution of T are considered. Upper 5% and 1% points are found for each of the approximations in the cases $m = 2, 3$, and 4 and some accuracy comparisons are made.

4.2 Approximations to the Distribution and Comparison of Percentage Points of T

In view of the known distribution of T_0^2 for the special cases mentioned in Section 4.1, a number of approximations to its distribution are proposed for the general situation. Of course we would like the approximation to be a simple one in terms of tabulated functions and to

satisfy as many known results as possible. Naturally, we would want any approximation to be sufficiently accurate for practical purposes. For these reasons, the following three approximations to the distribution of T are considered:

Approximation 1

$$T \stackrel{\cdot}{\sim} \alpha \chi_{\nu}^2(\gamma) \quad (4.2.1)$$

for some α , ν , and γ ;

Approximation 2

$$T \stackrel{\cdot}{\sim} \alpha F_{m\nu_1, \nu_2 - m + 1}(\gamma) \quad (4.2.2.)$$

for some α and γ ; and

Approximation 3

$$T \stackrel{\cdot}{\sim} \alpha F_{\nu, \nu_2 - m + 1}(\gamma) \quad (4.2.3)$$

for some α , ν , and γ .

Approximation 1 was chosen in view of the fact that the limiting distribution of T_0^2 is

$$\chi_{m\nu_1}^2 \left(\sum_{i=1}^m \lambda_i \right).$$

It was suspected that this approximation would be particularly suited for large ν_2 . Approximation 2 was chosen since in the cases $m = 1$ or $\nu_1 = 1$, the distribution of T

is given by (3.3.23) and (3.3.27), respectively. It was thought that approximation 2 would be fairly accurate overall. The third approximation was chosen for reasons similar to, and in hopes of improving, approximation 2.

In approximations 1 and 3, the first two null moments of the approximating distribution were equated to the first two null moments of T given in (3.3.32) and (3.3.33) in order to determine the unknown multipliers, α , and unknown degrees of freedom, v . In approximation 2, the first moment of the approximating distribution was equated to the first null moment of T given in (3.3.32) in order to determine the unknown multiplier, α . The first nonnull moment of each of the approximating distributions was then equated with the first nonnull moment of T given in (3.3.21) to obtain the unknown noncentrality parameters, γ . Derivation of these unknown multipliers, degrees of freedom, and noncentrality parameters follow.

Let u be a random variable such that

$$u \sim \alpha \chi_v^2(\gamma). \quad (4.2.4)$$

The first two null moments of u are known to be

$$E(u | \gamma=0) = \alpha v \quad (4.2.5)$$

and

$$E(u^2 | \gamma=0) = \alpha^2 v(v+2). \quad (4.2.6)$$

Equating these two expressions with (3.3.32) and (3.3.33) we obtain

$$\alpha = \frac{(v_2 - 1)(v_1 + v_2 - m - 1)}{(v_2 - m)(v_2 - m - 1)(v_2 - m - 3)} \quad (4.2.7)$$

and

$$v = \frac{mv_1(v_2 - m)(v_2 - m - 3)}{(v_2 - 1)(v_1 + v_2 - m - 1)}. \quad (4.2.8)$$

The first nonnull moment of u is given by

$$E(u|\gamma) = \alpha(v + \gamma). \quad (4.2.9)$$

Equating this with (3.3.21) and using (4.2.7) and (4.2.8) yields

$$\begin{aligned} \gamma &= \frac{(v_2 - m)(v_2 - m - 3)}{(v_2 - 1)(v_1 + v_2 - m - 1)} \sum_{i=1}^m \lambda_i \\ &= \frac{v}{mv_1} \sum_{i=1}^m \lambda_i, \end{aligned} \quad (4.2.10)$$

where the value of v is given by (4.2.8). Thus approximation 1 becomes

$$T \hat{\sim} \alpha \chi_v^2(\gamma), \quad (4.2.11)$$

where α , v , and γ are given by (4.2.7), (4.2.8) and (4.2.10) respectively.

If u is a random variable such that

$$u \sim \alpha F_{mv_1, v_2 - m + 1}(\gamma), \quad (4.2.12)$$

then the first null moment of u is

$$E(u|\gamma=0) = \alpha \frac{(v_2-m+1)}{(v_2-m-1)}. \quad (4.2.13)$$

The evaluation of α by equating (4.2.13) with (3.3.32) yields

$$\alpha = \frac{mv_1}{v_2-m+1}. \quad (4.2.14)$$

Using (3.3.21), (4.2.14) and the first nonnull moment of u ,

$$E(u|\gamma) = \alpha \frac{(mv_1+\gamma)(v_2-m+1)}{mv_1(v_2-m-1)}, \quad (4.2.15)$$

we obtain

$$\gamma = \sum_{i=1}^m \lambda_i. \quad (4.2.16)$$

Approximation 2 then becomes

$$T \hat{=} \frac{mv_1}{v_2-m+1} F_{mv_1, v_2-m-1} \left(\sum_{i=1}^m \lambda_i \right). \quad (4.2.17)$$

If u is a random variable with

$$u \sim \alpha F_{v, v_2-m+1}(\gamma), \quad (4.2.18)$$

the first two null moments and the first nonnull moment of u are respectively

$$E(u|\gamma=0) = \alpha \frac{(v_2-m+1)}{(v_2-m-1)}, \quad (4.2.19)$$

$$E(u^2 | \gamma=0) = \alpha^2 \frac{(v+2)(v_2-m+1)^2}{v(v_2-m-1)(v_2-m-3)}, \quad (4.1.20)$$

and

$$E(u | \gamma) = \alpha \frac{(v+\gamma)(v_2-m+1)}{v(v_2-m-1)}. \quad (4.2.21)$$

A procedure similar to that used for approximation 1 yields the following values for α , v , and γ :

$$\alpha = \frac{mv_1}{v_2-m+1}, \quad (4.2.22)$$

$$v = \frac{mv_1(v_2-m)}{v_1+v_2-mv_1-1}, \quad (4.2.23)$$

and

$$\begin{aligned} \gamma &= \frac{v_2-m}{v_1+v_2-mv_1-1} \sum_{i=1}^m \lambda_i \\ &= \frac{v}{mv_1} \sum_{i=1}^m \lambda_i. \end{aligned} \quad (4.2.24)$$

It should be pointed out that the value of v given in (4.2.23) may be negative when m and v_1 are large, relative to v_2 . We will agree to use approximation 3, therefore, only when the value of v is positive. That is, when $v_2 > (m-1)v_1+1$. Thus, approximation 3 becomes

$$T \stackrel{\circ}{=} \frac{mv_1}{v_2-m+1} F_{v, v_2-m+1} \left(\frac{v}{mv_1} \sum_{i=1}^m \lambda_i \right), \quad (4.2.25)$$

provided $v > 0$, where

$$v = \frac{mv_1(v_2-m)}{v_1+v_2-mv_1-1}. \quad (4.2.26)$$

Approximations 2 and 3 become exact in the cases $m = 1$ for all v_1 and v_2 ; $v_1 = 1$ for all m and v_2 ; and in the limit as v_2 goes to infinity. They are also invariant under the transformation

$$(m, v_1, v_2) \rightarrow (v_1, m, v_1+v_2-m).$$

Approximation 1 is also invariant under this transformation and becomes exact in the limit. It is obviously not exact, however, when $m = 1$ or $v_1 = 1$.

The following tables give the upper 5% and 1% points of T obtained from each of the three approximations considered in this chapter, as well as some accuracy comparisons. In the case $m = 2$, our approximate percentage points are compared with the exact percentage points obtained from the distribution derived by Hotelling [13] as well as those found using Pillai's approximation [29]. For $m = 3$ and 4, comparisons are made with Pillai's approximate percentage points and those given by Davis [7]. Hotelling's exact percentage points are reprinted from [30]. Percentage points for approximation 1 were obtained using Percentage Points of the χ^2 -distribution [28]. The percentage points for approximations 2 and 3, as well as

TABLE 4.1
Upper 5% points of T
(m = 2)

		v_2							
v_1		13	23	33	43	53	63	83	125
3	E	1.4508	.6807	.4429	.3280	.2603	.2158	.1607	.1064
	P	1.3966	.6683	.4385	.3261	.2596	.2156	.1610	.1069
	1	1.4994	.6861	.4446	.3286	.2606	.2159	.1607	.1064
	2	1.4979	.6952	.4524	.3341	.2648	.2192	.1631	.1078
	3	1.4473	.6820	.4447	.3297	.2619	.2172	.1619	.1073
5	E	2.1771	1.0071	.6517	.4812	.3813	.3157	.2348	.1552
	P	2.0893	.9849	.6424	.4763	.3784	.3138	.2340	.1550
	1	2.2284	1.0097	.6515	.4807	.3808	.3153	.2345	.1551
	2	2.2945	1.0439	.6703	.4924	.3889	.3212	.2381	.1569
	3	2.1741	1.0073	.6525	.4821	.3821	.3164	.2354	.1557
7	E	2.8802	1.3197	.8508	.6269	.4960	.4103	.3049	.2013
	P	2.7658	1.2907	.8384	.6202	.4920	.4076	.3035	.2008
	1	2.9324	1.3193	.8488	.6253	.4949	.4094	.3043	.2010
	2	3.0819	1.3829	.8817	.6454	.5084	.4192	.3101	.2038
	3	2.8828	1.3212	.8521	.6279	.4970	.4111	.3055	.2017
13	E	4.9409	2.2260	1.4243	1.0451	.8247	.6809	.5046	.3323
	P	4.7566	2.1806	1.4050	1.0346	.8182	.6765	.5023	.3314
	1	4.9920	2.2157	1.4169	1.0402	.8213	.6784	.5032	.3317
	2	5.4135	2.3783	1.4998	1.0903	.8549	.7024	.5173	.3383
	3	. . .	2.2342	1.4288	1.0480	.8268	.6825	.5056	.3329

TABLE 4.2
Upper 1% points of T
(m = 2)

		v_2							
v_1		13	23	33	43	53	63	83	123
3	E	2.2704	.9850	.6239	.4558	.3589	.2959	.2189	.1440
	P	2.0541	.9350	.6029	.4445	.3519	.2912	.2165	.1430
	1	2.1429	.9740	.6068	.4460	.3525	.2914	.2164	.1428
	2	2.4103	1.0250	.6438	.4675	.3666	.2820	.2222	.1456
	3	2.3025	.9953	.6293	.4593	.3613	.2977	.2201	.1447
5	E	3.2642	1.3893	.8739	.6361	.4998	.4114	.3039	.1995
	P	2.9624	1.3218	.8456	.6208	.4902	.4050	.3004	.1980
	1	3.0447	1.3287	.8467	.6209	.4900	.4947	.3001	.1978
	2	3.5800	1.4807	.9167	.6609	.5159	.4228	.3104	.2025
	3	3.3335	1.4100	.8840	.6422	.5038	.4143	.3056	.2003
7	E	4.2220	1.7736	1.1099	.8058	.6320	.5197	.3833	.2513
	P	3.8445	1.6922	1.0764	.7878	.6208	.5121	.3792	.2495
	1	3.9124	1.6911	1.0735	.7855	.6191	.5108	.3783	.2491
	2	4.7468	1.9223	1.1796	.8461	.6583	.5382	.3939	.2561
	3	4.3327	1.8065	1.1258	.8152	.6382	.5241	.3859	.2525
13	E	7.0198	2.8804	1.7842	1.2879	1.0065	.8255	.6069	.3964
	P	6.4298	2.7618	1.7373	1.2634	.9915	.8155	.6015	.3942
	1	6.4443	2.7339	1.7208	1.2530	.9844	.8104	.5985	.3928
	2	8.1696	3.2149	1.9432	1.3809	1.0675	.8687	.6317	.4078
	3	. . .	2.9509	1.8182	1.3079	1.0196	.8349	.6123	.3989

TABLE 4.3
Upper 5% points of T
(m = 3)

		ν_2						
ν_1		30	40	50	60	80	100	200
5	D	1.0385	.7348	.5682	.4631	.3380	.2661	.1289
	P	1.0097	.7201	.5595	.4574	.3351	.2644	.1286
	1	1.0361	.7329	.5668	.4621	.3374	.2657	.1288
	2	1.0936	.7655	.5897	.4768	.3458	.2712	.1303
	3	1.0407	.7364	.5695	.4642	.3388	.2667	.1292
10	D	1.8477	1.3020	1.0045	.8174	.5954	.4682	.2263
	P	1.8022	1.2791	.9908	.8084	.5907	.4653	.2257
	1	1.8379	1.2960	1.0005	.8146	.5938	.4672	.2260
	2	2.0026	1.3892	1.0605	.8564	.6175	.4824	.2299
	3	1.8579	1.3081	1.0085	.8203	.5972	.4694	.2266
15	D	2.6327	1.8503	1.4251	1.1584	.8425	.6619	.3193
	P	2.5731	1.8206	1.4075	1.1468	.8365	.6583	.3185
	1	2.6157	1.8402	1.4181	1.1531	.8394	.6599	.3189
	2	2.8997	2.0022	1.5234	1.2273	.8819	.6874	.3259
	3	. . .	1.8616	1.4325	1.1636	.8456	.6639	.3199

TABLE 4.4
Upper 1% points of T
(m = 3)

		ν_2						
ν_1		30	40	50	60	80	100	200
5	D	1.3504	.9389	.7189	.5824	.4219	.3305	.1589
	P	1.2773	.9024	.6973	.5680	.4143	.3261	.1578
	1	1.3041	.9152	.7047	.5729	.4168	.3275	.1582
	2	1.4749	1.0061	.7611	.6112	.4379	.3408	.1614
	3	1.3785	.9544	.7288	.5892	.4257	.3332	.1595
10	D	2.2913	1.5848	1.2101	.9783	.7070	.5534	.2652
	P	2.1863	1.5341	1.1805	.9590	.6970	.5474	.2638
	1	2.1853	1.5442	1.1859	.9623	.6986	.5482	.2640
	2	2.6151	1.7620	1.3219	1.0053	.7499	.5808	.2719
	3	2.3590	1.6221	1.2337	.9945	.7161	.5592	.2666
15	D	3.1977	2.2039	1.6791	1.3555	.9778	.7645	.3654
	P	3.0638	2.1406	1.6427	1.3320	.9658	.7572	.3638
	1	3.0826	2.1469	1.6447	1.3322	.9653	.7568	.3637
	2	3.7372	2.5019	1.8684	1.4865	1.0514	.8116	.3773
	3	. . .	2.2629	1.7163	1.3811	.9920	.7736	.3677

TABLE 4.5
Upper 5% points of T
(m = 4)

		v_2						
v_1		30	40	50	60	80	100	200
5	D	1.3638	.9524	.7313	.5934	.4308	.3381	.1629
	P	1.3140	.9274	.7165	.5836	.4257	.3350	.1622
	1	1.3589	.9491	.7290	.5918	.4299	.3375	.1627
	2	1.4621	1.0062	.7653	.6168	.4439	.3465	.1650
	3	1.3689	.9556	.7335	.5951	.4319	.3389	.1631
10	D	2.4501	1.7053	1.3070	1.0593	.7679	.6022	.2896
	P	2.3732	1.6676	1.2847	1.0447	.7603	.5975	.2886
	1	2.4357	1.6971	1.3017	1.0553	.7654	.6005	.2891
	2	2.7213	1.8562	1.4033	1.1261	.8056	.6264	.2957
	3	. . .	1.7166	1.3143	1.0644	.7709	.6041	.2901
15	D	3.5074	2.4358	1.8643	1.5095	1.0930	.8565	.4113
	P	3.4069	2.3871	1.8359	1.4910	1.0834	.8507	.4100
	1	3.4830	2.4212	1.8560	1.5034	1.0896	.8545	.4107
	2	3.9696	2.6961	2.0318	1.6266	1.1597	.8996	.4223
	3	. . .	. . .	1.8773	1.5186	1.0982	.8599	.4121

TABLE 4.6
Upper 1% points of T
(m = 4)

		ν_2						
ν_1		30	40	50	60	80	100	200
5	D	1.7348	1.1892	.9040	.7289	.5253	.4106	.1962
	P	1.6178	1.1324	.8706	.7071	.5139	.4035	.1946
	1	1.6745	1.1595	.8865	.7174	.5193	.4069	.1953
	2	1.9498	1.3036	.9749	.7773	.5519	.4273	.2003
	3	1.7819	1.2147	.9200	.7399	.5314	.4144	.1972
10	D	2.9810	2.0356	1.5441	1.2434	.8946	.6985	.3332
	P	2.8126	1.9564	1.4985	1.2139	.8795	.6894	.3312
	1	2.8785	1.9865	1.5155	1.2244	.8843	.6920	.3316
	2	3.5341	2.3341	1.7301	1.3716	.9657	.7437	.3444
	3	. . .	2.0957	1.5816	1.2690	.9087	.7075	.3354
15	D	4.1858	2.8501	2.1581	1.7359	1.2471	.9729	.4632
	P	3.9706	2.7510	2.1019	1.6999	1.2289	.9619	.4608
	1	4.0413	2.7801	2.1188	1.7097	1.2335	.9648	.4613
	2	5.1018	3.3493	2.4732	1.9531	1.3686	1.0506	.4827
	3	. . .	. . .	2.2170	1.7760	1.2692	.9869	.4666

Pillai's approximation, were obtained from Tables of the Incomplete Beta Distribution [28]. There are five entries in each cell. For $m = 2$, these correspond to the exact percentage points and the approximate percentage points found from Pillai's approximation and approximations 1, 2 and 3 of this chapter. These are designated by E, P, 1, 2, and 3 respectively. For $m = 3$ and 4 the exact percentage points are replaced by those given by Davis which are designated by D. Davis' percentage points are then used as a standard by which to compare approximations. The omitted percentage points for approximation 3 are those which occurred in cases where $v_2 \leq (m - 1) v_1 + 1$.

4.3 Summary

A comparison of the preceding tables reveals the following generality for all approximations. The 5% points are uniformly more accurate than the corresponding 1% points, and the accuracy of both the 5% and 1% points increased with v_2 , while decreasing with both v_1 and m .

Some unexpected results are also revealed. The 5% and 1% points of T computed from the χ^2 approximation (approximation 1) are surprisingly close to the exact values for $m = 2$ and the values given by Davis in the cases $m = 3$ and 4, even for small v_2 . In comparing this approximation with Pillai's, the χ^2 approximation is

generally more accurate for both the 5% and 1% points. Some exceptions occur in the 1% points when both v_1 and v_2 are large.

Although it was hoped that approximation 2 would be relatively accurate overall, the results were rather disappointing. This approximation is, without exception, less accurate than the others. It is, however, very simple, and fairly accurate for large values of v_2 .

The results of approximation 3 were very satisfying. Both the 5% and 1% points for this approximation were generally more accurate than Pillai's, and markedly so for small v_2 . The decrease in accuracy for increasing v_1 seems to be less for approximation 3 than for the other approximations. The fact that this approximation cannot be applied when $v_2 \leq (m - 1) v_1 + 1$ should not be too disturbing, since in most practical situations we would expect v_2 to be large, relative to v_1 .

Although the comparison of percentage points is only for the cases $m = 2, 3$ and 4 and selected values of v_1 and v_2 , it is suspected that the relative accuracy of the approximations would be maintained for general m , v_1 and v_2 . It is therefore suggested that the distribution of T be approximated by

$$T \hat{=} \frac{mv_1}{v_2 - m + 1} F_{v, v_2 - m + 1} \left(\frac{v}{mv_1} \sum_{i=1}^m \lambda_i \right) \quad (4.3.1)$$

when $v_2 > (m - 1) v_1 + 1$. In this case the value of v is given by

$$v = \frac{mv_1(v_2 - m)}{v_2 - (m-1)v_1 + 1}. \quad (4.3.2)$$

If the situation should arise when $v_2 \leq (m - 1) v_1 + 1$, the χ^2 approximation is suggested.

$$T \hat{=} \alpha \chi_v^2 \left(\frac{v}{mv_1} \sum_{i=1}^m \lambda_i \right) \quad (4.3.3)$$

in which

$$\alpha = \frac{(v_2 - 1)(v_1 + v_2 - m - 1)}{(v_2 - m)(v_2 - m - 1)(v_2 - m - 3)} \quad (4.3.4)$$

and

$$v = \frac{mv_1(v_2 - m)(v_2 - m - 3)}{(v_2 - 1)(v_1 + v_2 - m - 1)}. \quad (4.3.5)$$

CHAPTER V

THE POWER OF HOTELLING'S GENERALIZED T_0^2

5.1 Introduction

In Chapter I it was noted that the power function of T_0^2 , or equivalently of T ,

$$\beta(\Lambda; m, v_1, v_2; t) = P(T > t | \Lambda), \quad (5.1.1)$$

where Λ is the $m \times m$ diagonal matrix of elements $\lambda_1, \dots, \lambda_m$ defined in (1.1.23), is a function only of m , v_1 , v_2 and $\lambda_1, \dots, \lambda_m$. This power function is known to be an increasing function of each λ_i [5], and T_0^2 is known to be an admissible test of the general linear hypothesis [10]. Aside from these properties, little is known concerning the behavior of the power function of T_0^2 in the general case.

Pillai and Jayachandran [30] have compared the powers of four test criteria (see (1.1.24), (1.1.25), (1.1.26) and (1.1.27)) for testing the general linear hypothesis in the case $m = 2$ for selected values of v_1 and v_2 . On the basis of these power comparisons, Pillai and Jayachandran noted some apparent properties of these criteria, summarized below.

1. Roy's largest root criterion seems to have the smallest power overall of the four test criteria considered.

2. The power of Pillai's criteria is generally somewhat greater than that of the other three criteria. Some exceptions were noted, however, with respect to the likelihood ratio test and T_0^2 for large v_2 when λ_1 and λ_2 are not nearly equal.
3. T_0^2 seems to have greater power than the likelihood ratio for large v_2 except when λ_1 and λ_2 are nearly equal.
4. None of the four criteria considered seem to have a monotonic power function in the sum $\lambda_1 + \lambda_2$. For fixed $\lambda_1 + \lambda_2$, as λ_1 and λ_2 tend to be equal, the powers of Pillai's criteria and the likelihood ratio test increase, while the powers of T_0^2 and Roy's largest root decrease.

With reference to remark 4 above, it should be pointed out that for any two sets of alternatives, Λ_1 and Λ_2 , where Λ_i : $i = 1, 2$, is an $m \times m$ diagonal matrix with diagonal elements λ_{ij} : $j = 1, \dots, m$, such that

$$\text{tr } \Lambda_1 = \text{tr } \Lambda_2, \quad (5.1.2)$$

T_0^2 cannot have greater power against one of these alternatives for all levels of significance. In order to see this, consider

$$\begin{aligned} \int_0^\infty P(T > t | \Lambda_1) dt &= E(T | \Lambda_1) \\ &= \frac{mv_1 + \text{tr } \Lambda_1}{v_2 - m - 1}. \end{aligned} \quad (5.1.3)$$

Equations (5.1.2) and (5.1.3) imply that

$$\int_0^\infty P(T > t | \Lambda_1) dt = \int_0^\infty P(T > t | \Lambda_2) dt, \quad (5.1.4)$$

so that it cannot be true that

$$P(T > t | \Lambda_i) > P(T > t | \Lambda_j), \quad (5.1.5)$$

for all $t \geq 0$.

Rather than give further power comparisons of various test criteria, we present here some integral bounds on the power function of T_0^2 . These are based on a result due to Wald [35] and Hsu [16], and the properties of completely monotone functions [9]. In addition, we offer an alternative proof of the result obtained by Das Gupta, Mudholkar and Anderson.

5.2 *An Upper Bound on the Integral of the Power Function of T_0^2*

In 1942, Wald [35] and in 1945, Hsu [16] proved the following result in connection with the analysis of variance. Suppose we are given $z_1, \dots, z_p$ and r^2 with the variates independently distributed such that

$$z_i \sim N(\xi_i, \sigma^2): \quad i = 1, \dots, p \quad (5.2.1)$$

and

$$r^2 \sim \sigma^2 \chi_n^2(0). \quad (5.2.2)$$

It is desired to test the hypothesis that

$$\xi_1 = \dots = \xi_p = 0 \quad (5.2.3)$$

against the general alternative. Let ω^* be the critical region of the form

$$\omega^* = \left\{ (z_1, \dots, z_p, r^2) : \sum_{i=1}^p z_i^2 / r^2 > c \right\}, \quad (5.2.4)$$

where c is chosen so that ω^* is of some desired size, α .

The power function of ω^* ,

$$\beta_0(\eta; p, n) = P \left[\sum_{i=1}^p z_i^2 / r^2 > c \mid \eta \right], \quad (5.2.5)$$

depends only upon p , n and the parameter

$$\eta = \frac{1}{\sigma^2} \sum_{i=1}^p \xi_i^2. \quad (5.2.6)$$

Suppose ω is any other critical region of the same size as ω^* , and, with

$$\underline{\xi} = (\xi_1, \dots, \xi_p)', \quad (5.2.7)$$

denote its power function by

$$\beta_\omega(\underline{\xi}; p, n). \quad (5.2.8)$$

For fixed η , let $S(\underline{\xi}, \sigma, \eta)$ denote the sphere defined by (5.2.6) so that

$$A(\underline{\xi}, \sigma, \eta) = \int_{S(\underline{\xi}, \sigma, \eta)} d\underline{\xi} \quad (5.2.9)$$

is the surface area of this sphere. Then

$$\beta_0(n;p,n) \geq \frac{1}{A(\underline{\xi},\sigma,n)} \int_{S(\underline{\xi},\sigma,n)} \beta_\omega(\underline{\xi};p,n) d\underline{\xi}. \quad (5.2.10)$$

We use this result in order to obtain an upper bound on the integral of the power function of T_0^2 . We have an $m \times v_1$ matrix, Y , and an $m \times m$ matrix, S_2 , where Y and S_2 are independently distributed with

$$Y \sim N_{m \times v_1}(\underline{E}, I \otimes I) \quad (5.2.11)$$

and

$$S_2 \sim W_m(I, v_2, (0)) \quad (5.2.12)$$

(see (3.1.6) and (3.1.10)). It is desired to test the hypothesis

$$H_0: \underline{E} = (0)$$

against

$$(5.2.13)$$

$$H_a: \underline{E} \neq (0).$$

With S_2 distributed as in (5.2.12) we may write (see [1] or [32])

$$S_2 = RR', \quad (5.2.14)$$

where R is the unique $m \times m$ lower triangular matrix defined by (5.2.14) such that

$$r_{ii} \sim \chi_{v_2-i+1}(0): \quad i = 1, \dots, m \quad (5.2.15)$$

and

$$r_{ij} \sim N(0,1): \quad i > j \quad (5.2.16)$$

with all variates independent. Then

$$\begin{aligned} \text{tr } S_2 &= \text{tr } RR' \\ &= \sum_{i=1}^m \sum_{j=1}^i r_{ij}^2 \\ &\sim \chi_{mv_2}^2(0). \end{aligned} \quad (5.2.17)$$

From (5.2.11)

$$y_{ij} \sim N(\xi_{ij}, 1): \quad i = 1, \dots, m; \quad j = 1, \dots, v_1. \quad (5.2.18)$$

Let ω^* be the critical region of size α for testing (5.2.13) of the form

$$\omega^* = \left\{ (Y, S_2): \sum_{i=1}^m \sum_{j=1}^{v_1} y_{ij}^2 / \text{tr } S_2 > c \right\}. \quad (5.2.19)$$

The power function of ω^* ,

$$\beta_0(\eta; m, v_1, v_2; c), \quad (5.2.20)$$

depends only upon m , v_1 , v_2 and the parameter

$$\begin{aligned} \eta &= \sum_{i=1}^m \sum_{j=1}^{v_1} \xi_{ij}^2 \\ &= \text{tr } \Xi \Xi' \\ &= \text{tr } \Lambda \end{aligned} \quad (5.2.21)$$

(see equations (3.1.7) and (3.3.4)). Then from the result due to Wald and Hsu (5.2.10)

$$\begin{aligned} \beta_0(n; m, v_1, v_2; c) &\geq \frac{1}{A(\Lambda; m, v_1, \Xi)} \int_{S(\Lambda; m, v_1, \Xi)} \\ &\times \beta(\Lambda; m, v_1, v_2; t) d\Xi, \end{aligned} \quad (5.2.22)$$

where $S(\Lambda; m, v_1, \Xi)$ denotes the sphere defined by (5.2.21) and

$$A(\Lambda; m, v_1, \Xi) = \int_{S(\Lambda; m, v_1, \Xi)} d\Xi, \quad (5.2.23)$$

its surface area. From (5.2.11), (5.2.17) and (5.2.19)

$$\beta_0(n; m, v_1, v_2; c) = P\{\chi_{mv_1}^2(\text{tr } \Lambda) / \chi_{mv_2}^2(0) > c\}. \quad (5.2.24)$$

Thus

$$\begin{aligned} &P\{\chi_{mv_1}^2(\text{tr } \Lambda) / \chi_{mv_2}^2(0) > c\} \\ &\geq \frac{1}{A(\Lambda; m, v_1, \Xi)} \int_{S(\Lambda; m, v_1, \Xi)} \beta(\Lambda; m, v_1, v_2; t) d\Xi \end{aligned} \quad (5.2.25)$$

is an upper bound on the integral of the power function of T_0^2 . It is noted, of course, that c and t are chosen to satisfy

$$\begin{aligned} &P\{\chi_{mv_1}^2(0) / \chi_{mv_2}^2(0) > c\} \\ &= \beta\{(0); m, v_1, v_2; t\} \\ &= \alpha \end{aligned} \quad (5.2.26)$$

for some level of significance, α .

5.3 *A Lower Bound on the Integral of the
Power Function of T_0^2*

The lower bound on the integral of $\beta(\Lambda; m, v_1, v_2; t)$ is based on the properties of completely monotone functions as described by Feller [9]. For reference, we note some of the basic properties of these functions.

A function, $\phi(\theta)$, $0 < \theta < \infty$, is said to be completely monotone (θ) if $\phi(\theta)$ possesses derivatives of all orders, $\phi^{(n)}(\theta)$, and

$$(-1)^n \phi^{(n)}(\theta) \geq 0; \quad 0 < \theta < \infty. \quad (5.3.1)$$

Some properties of completely monotone functions, which follow directly from the above definition, are listed below.

1. If $\phi(\theta)$ is completely monotone (θ), then $(-1)^n \phi^{(n)}(\theta)$ is completely monotone (θ).
2. If $\phi(\theta)$ and $\psi(\theta)$ are completely monotone (θ), then

$$R_1(\theta) = \phi(\theta) + \psi(\theta) \quad (5.3.2)$$

and

$$R_2(\theta) = \phi(\theta) \psi(\theta) \quad (5.3.3)$$

are completely monotone (θ).

3. If $\phi(\theta)$ is completely monotone (θ) and $\psi(\theta)$ is a positive function whose first derivative is completely monotone (θ), then

$$R_3(\theta) = \Phi[\Psi(\theta)] \quad (5.3.4)$$

is completely monotone (θ) .

The main property of completely monotone functions is the following theorem, proved in Feller [9].

Theorem: A function $\Phi(\theta)$ on $0 < \theta < \infty$ is completely monotone (θ) if and only if

$$\Phi(\theta) = \int_0^{\infty} e^{-\theta x} F(dx) \quad (5.3.5)$$

where $F(x)$ is a (possibly infinite) measure on $[0, \infty]$.

It follows that if $\Phi(\theta)$ can be written in the form

$$\Phi(\theta) = \int_0^{\infty} h(x) e^{-\theta x} dx, \quad (5.3.6)$$

then $\Phi(\theta)$ is completely monotone (θ) if and only if $h(x)$ is nonnegative on $[0, \infty]$.

Recall the representation of T given in (3.2.34)

$$T = \sum_{i=1}^m \frac{u_i^2}{d_i}, \quad (5.3.7)$$

where, conditional on the $m \times m$ random orthogonal matrix, K (see equations (3.2.25), (3.2.29), (3.3.30) and (3.3.31)) $u_1^2, \dots, u_m^2$ are independently distributed with

$$u_i^2 | K \sim \chi_{\nu_1}^2(\gamma_i): \quad i = 1, \dots, m \quad (5.3.8)$$

and

$$\begin{aligned} \gamma_i &= (K \Lambda K')_{ii} \\ &= \sum_{j=1}^m k_{ij}^2 \lambda_j. \end{aligned} \quad (5.3.9)$$

The variates $d_1, \dots, d_m$ are the latent roots of S_2 . For convenience, write

$$\alpha_i = \frac{1}{d_i}: \quad i = 1, \dots, m. \quad (5.3.10)$$

Thus, using (5.3.7), (5.3.8), (5.3.9) and (5.3.10), for fixed $\underline{\alpha}' = (\alpha_1, \dots, \alpha_m)$ and K , we have

$$T | \underline{\alpha}, K \sim \sum_{i=1}^m \alpha_i \chi_{v_1}^2(\gamma_i). \quad (5.3.11)$$

We wish to consider the function

$$\Phi(\theta) = \int_0^\infty \left[\int_s^\infty P(T > t | \underline{\alpha}, K, \Lambda) dt \right] e^{-\theta s} ds. \quad (5.3.12)$$

The inside integral of (5.3.12) is nonnegative and hence from (5.3.6), $\Phi(\theta)$ is completely monotone (θ). Integration of (5.3.12) along with (5.3.11) yields

$$\begin{aligned} \Phi(\theta) &= \frac{1}{\theta} \sum_{i=1}^m \alpha_i (v_1 + \gamma_i) - \frac{1}{\theta^2} \\ &\quad + \frac{1}{\theta^2} \prod_{i=1}^m (1 + 2\theta\alpha_i)^{-v_1/2} \exp \left[-\frac{\theta\alpha_i\gamma_i}{1 + 2\theta\alpha_i} \right]. \end{aligned} \quad (5.3.13)$$

Writing

$$g_i(\delta) = (1 + 2\delta)^{-v_1/2} \exp \left[-\frac{\delta\gamma_i}{1 + 2\delta} \right], \quad (5.3.14)$$

we have

$$\Phi(\theta) = \frac{1}{\theta} \sum_{i=1}^m \alpha_i (v_1 + \gamma_i) - \frac{1}{\theta^2} + \frac{1}{\theta^2} \prod_{i=1}^m g_i(\theta \alpha_i). \quad (5.3.15)$$

It should be noted that $g_i(\delta)$ defined in (5.3.14) is the Laplace transform of a random variable distributed as $\chi_{v_1}^2(\gamma_i)$. Hence, from (5.3.6), $g_i(\delta)$ is completely monotone (δ).

A function, $f(x_1, \dots, x_m)$, is convex in $x_1, \dots, x_m$ if and only if its Hessian, the $m \times m$ matrix, H , of second-order partial derivatives,

$$h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}: \quad i = 1, \dots, m; \quad j = 1, \dots, m, \quad (5.3.16)$$

is nonnegative definite. Consider the second-order derivatives of $\Phi(\theta)$ with respect to $\alpha_1, \dots, \alpha_m$. Direct differentiation shows that

$$\frac{\partial^2 \Phi(\theta)}{\partial \alpha_i^2} = \frac{(v_1 + \gamma_i)(v_1 + \gamma_i + 2)}{(1 + 2\alpha_i \theta)^2} \prod_{j=1}^m g_j(\theta \alpha_j) \quad (5.3.17)$$

and

$$\frac{\partial^2 \Phi(\theta)}{\partial \alpha_i \partial \alpha_j} = \frac{(v_1 + \gamma_i)(v_1 + \gamma_j)}{(1 + 2\theta \alpha_i)(1 + 2\theta \alpha_j)} \prod_{k=1}^m g_k(\theta \alpha_k). \quad (5.3.18)$$

Thus the Hessian becomes

$$H = H_1 \prod_{i=1}^m g_i(\theta \alpha_i), \quad (5.3.19)$$

where H_1 is the $m \times m$ matrix

$$H_1 = \begin{bmatrix} \frac{(\nu_1 + \gamma_1)(\nu_1 + \gamma_1 + 2)}{(1 + 2\theta\alpha_1)^2} & \cdot & \cdot & \cdot & \frac{(\nu_1 + \gamma_1)(\nu_1 + \gamma_m)}{(1 + 2\theta\alpha_1)(1 + 2\theta\alpha_m)} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \frac{(\nu_1 + \gamma_1)(\nu_1 + \gamma_m)}{(1 + 2\theta\alpha_1)(1 + 2\theta\alpha_m)} & \cdot & \cdot & \cdot & \frac{(\nu_1 + \gamma_m)(\nu_1 + \gamma_m + 2)}{(1 + 2\theta\alpha_m)^2} \end{bmatrix}. \quad (5.3.20)$$

The matrix H_1 can be written as the sum

$$H_1 = H_2 + H_3, \quad (5.3.21)$$

where

$$H_2 = \begin{bmatrix} \frac{(\nu_1 + \gamma_1)}{(1 + 2\theta\alpha_1)} \\ \cdot \\ \cdot \\ \cdot \\ \frac{(\nu_1 + \gamma_m)}{(1 + 2\theta\alpha_m)} \end{bmatrix} \begin{bmatrix} \frac{(\nu_1 + \gamma_1)}{(1 + 2\theta\alpha_1)}, & \cdot & \cdot & \cdot, & \frac{(\nu_1 + \gamma_m)}{(1 + 2\theta\alpha_m)} \end{bmatrix} \quad (5.3.22)$$

and

$$H_3 = \begin{bmatrix} \frac{2(\nu_1 + \gamma_1)}{(1 + 2\theta\alpha_1)^2} & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & \cdot & \cdot & \cdot & \frac{2(\nu_1 + \gamma_m)}{(1 + 2\theta\alpha_m)^2} \end{bmatrix}. \quad (5.3.23)$$

Since H_2 is nonnegative definite and H_3 is positive definite, H_1 , and therefore the Hessian H , is positive definite.

Using the definition of convexity, $\Phi(\theta)$ given by (5.3.12) or (5.3.15), is convex in $\alpha_1, \dots, \alpha_m$.

Jensen's inequality states that

$$E\{f(x_1, \dots, x_m)\} \geq f\{E(x_1), \dots, E(x_m)\} \quad (5.3.24)$$

whenever $f(x_1, \dots, x_m)$ is convex in $x_1, \dots, x_m$, provided of course that these expectations exist. From (5.3.10) and (3.3.18)

$$E(\alpha_i) = \frac{1}{v_2 - m - 1}: \quad i = 1, \dots, m. \quad (5.3.25)$$

Replacing α_i by its expectation in (5.3.15), denoting this new function by $\Phi_*(\theta)$, and applying Jensen's inequality to $\Phi(\theta)$, we have

$$\frac{E\{\Phi(\theta)\}}{\alpha} - \Phi_*(\theta) \geq 0, \quad (5.3.26)$$

where

$$\begin{aligned} \Phi_*(\theta) &= \int_0^\infty \left[\int_0^\infty P \left[E(T) > t \mid K, \Lambda \right] dt \right] e^{-\theta s} ds \\ &= \frac{mv_1 + \sum_{i=1}^m \gamma_i}{\theta(v_2 - m - 1)} - \frac{1}{\theta^2} + \frac{1}{\theta^2} \prod_{i=1}^m g_i \left(\frac{\theta}{v_2 - m - 1} \right). \end{aligned} \quad (5.3.27)$$

Letting

$$\Psi(\theta) = \underline{E}_{\underline{\alpha}}(\Phi(\theta)) - \Phi_{*}(\theta), \quad (5.3.28)$$

we wish to show that $\Psi(\theta)$ is completely monotone (θ) . We need to show that

$$(-1)^{n_{\Psi}(n)}(\theta) \geq 0: \quad n = 1, 2, \dots. \quad (5.3.29)$$

Now

$$\begin{aligned} (-1)^{n_{\Psi}(n)}(\theta) &= (-1)^n \frac{d^n}{d\theta^n} \left\{ \underline{E}_{\underline{\alpha}}(\Phi(\theta)) - \Phi_{*}(\theta) \right\} \\ &= \underline{E}_{\underline{\alpha}} \left\{ (-1)^{n_{\Phi}(n)}(\theta) \right\} - (-1)^{n_{\Phi_{*}}(n)}(\theta). \end{aligned} \quad (5.3.30)$$

Since $(-1)^{n_{\Phi_{*}}(n)}(\theta)$ is simply $(-1)^{n_{\Phi}(n)}(\theta)$ with $\alpha_1, \dots, \alpha_m$ replaced by their expectations, (5.3.29) will be proved if we can show that $(-1)^{n_{\Phi}(n)}(\theta)$ is convex in $\alpha_1, \dots, \alpha_m$ and applying Jensen's inequality. Consider

$$\frac{\partial^2}{\partial \alpha_i^2} \left\{ (-1)^{n_{\Phi}(n)}(\theta) \right\} = (-1)^n \frac{d^n}{d\theta^n} \frac{\partial^2 \Phi(\theta)}{\partial \alpha_i^2} \quad (5.3.31)$$

and

$$\frac{\partial^2}{\partial \alpha_i \partial \alpha_j} \left\{ (-1)^{n_{\Phi}(n)}(\theta) \right\} = (-1)^n \left[\frac{d^n}{d\theta^n} \frac{\partial^2 \Phi(\theta)}{\partial \alpha_i \partial \alpha_j} \right]. \quad (5.3.32)$$

From (5.3.19), (5.3.31) and (5.3.32), the Hessian of $(-1)^{n_{\Phi}(n)}(\theta)$ is

$$H^{*} = (-1)^n \frac{d^n}{d\theta^n} H$$

$$\begin{aligned}
 &= (-1)^n \frac{d^n}{d\theta^n} \left(H_1 \prod_{i=1}^m g_i(\theta \alpha_i) \right) \\
 &= \prod_{i=1}^m g_i(\theta \alpha_i) \left((-1)^n \frac{d^n}{d\theta^n} H_1 \right) + H_1 \left((-1)^n \frac{d^n}{d\theta^n} \prod_{i=1}^m g_i(\theta \alpha_i) \right),
 \end{aligned}
 \tag{5.3.33}$$

where H_1 is the positive definite matrix of (5.3.20). From (5.3.21), (5.3.22) and (5.3.23),

$$\begin{aligned}
 (-1)^n \frac{d^n}{d\theta^n} H_1 &= (-1)^n \frac{d^n}{d\theta^n} (H_2 + H_3) \\
 &= H_4 + H_4' + H_5,
 \end{aligned}
 \tag{5.3.34}$$

where

$$H_4 = \begin{bmatrix} \frac{(v_1 + \gamma_1) n! (2\alpha_1)^n}{(1 + 2\theta\alpha_1)^{n+1}} \\ \vdots \\ \frac{(v_1 + \gamma_m) n! (2\alpha_m)^n}{(1 + 2\theta\alpha_m)^{n+1}} \end{bmatrix} \begin{bmatrix} \frac{(v_1 + \gamma_1)}{(1 + 2\theta\alpha_1)}, \dots, \frac{(v_1 + \gamma_m)}{(1 + 2\theta\alpha_m)} \end{bmatrix} \tag{5.3.35}$$

and

$$H_5 = \begin{bmatrix} \frac{2(\nu_1 + \gamma_1)(n+1)!(2\alpha_1)^n}{(1+2\theta\alpha_1)^{n+2}} & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & \frac{2(\nu_1 + \gamma_m)(n+1)!(2\alpha_m)^n}{(1+2\theta\alpha_1)^{n+2}} \end{bmatrix}, \quad (5.3.36)$$

so that $(-1)^n d^n/d\theta^n H_1$ is positive definite. Recall that $g_i(\delta)$ is defined in (5.3.14) is completely monotone (δ) . The function

$$h_i(\theta) = \theta\alpha_i \quad (5.3.37)$$

is a positive function whose first derivative is completely monotone (θ) . Using property 3 of completely monotone functions (5.3.4), we see that $g_i(\theta\alpha_i)$ is completely monotone (θ) . Repeated application of property 2 (5.3.3) implies that

$$G(\theta) = \prod_{i=1}^m g_i(\theta\alpha_i) \quad (5.3.38)$$

is completely monotone (θ) . Thus

$$(-1)^n \frac{d^n}{d\theta^n} G(\theta) \geq 0: \quad n = 1, 2, \dots, \quad (5.3.39)$$

and hence, from (5.3.35) the Hessian, H^* , of $(-1)^{n_\Phi} \Phi^{(n)}(\theta)$

is nonnegative definite. This shows that $(-1)^{n_{\Phi}(n)}(\theta)$ is convex in $\alpha_1, \dots, \alpha_m$, which in turn shows that

$$\Psi(\theta) = \underline{E}_{\underline{\alpha}}(\Phi(\theta)) - \Phi_*(\theta) \quad (5.3.40)$$

is completely monotone (θ) .

The definition of $\Phi(\theta)$ is given in (5.3.12) and our definition of $\Phi_*(\theta)$ given in (5.3.27) yield

$$\begin{aligned} \Psi(\theta) = & \int_0^\infty \left[\int_s^\infty \underline{E}_{\underline{\alpha}}(P(T > t | \alpha, K, \Lambda)) dt \right. \\ & \left. - \int_s^\infty P(\underline{E}_{\underline{\alpha}}(T) > t | K, \Lambda) dt \right] e^{-\theta s} ds. \end{aligned} \quad (5.3.41)$$

Equation (5.3.6) and the complete monotonicity of $\Psi(\theta)$ imply that

$$\begin{aligned} & \int_s^\infty \underline{E}_{\underline{\alpha}}(P(T > t | \underline{\alpha}, K, \Lambda)) dt \\ & \geq \int_s^\infty P(\underline{E}_{\underline{\alpha}}(T) > t | K, \Lambda) dt, \end{aligned} \quad (5.3.42)$$

for all s , $0 \leq s < \infty$. From (5.3.7), (5.3.8), (5.3.9) (5.3.10) and (5.3.25) and since K is orthogonal,

$$P(\underline{E}_{\underline{\alpha}}(T) > t | K, \Lambda) = P\left[\underline{E}_{\underline{\alpha}}\left[\sum_{i=1}^m \alpha_i u_i^2\right] > t | K, \Lambda\right]$$

$$\begin{aligned}
 &= P\left(\frac{1}{v_2 - m - 1} \chi_{mv_1}^2(\text{tr } \Lambda) > t | K\right) \\
 &= P\left(\chi_{mv_1}^2(\text{tr } \Lambda) > (v_2 - m - 1) t\right) \quad (5.3.43)
 \end{aligned}$$

The last equation follows from the fact that the one immediately preceding this does not depend on K . Equations (5.3.42) and (5.3.43) then imply that

$$\begin{aligned}
 &\int_s^\infty \underline{E}_\alpha(P(T > t | \underline{\alpha}, K, \Lambda)) dt \\
 &\geq \int_s^\infty P(\chi_{mv_1}^2(\text{tr } \Lambda) > (v_2 - m - 1) t) dt, \quad (5.3.44)
 \end{aligned}$$

for all s , $0 \leq s < \infty$. Since (5.3.44) is true for all orthogonal K , it is true unconditionally on K . This, along with the fact that

$$\underline{E}_\alpha(P(T > t | \underline{\alpha}, \Lambda)) = P(T > t | \Lambda), \quad (5.3.45)$$

yields a lower bound on the integral of the power function of T_0^2 ,

$$\begin{aligned}
 &\int_s^\infty \beta(\Lambda; m, v_1, v_2; t) dt \\
 &\geq \int_s^\infty P(\chi_{mv_1}^2(\text{tr } \Lambda) > (v_2 - m - 1)t) dt, \quad (5.3.46)
 \end{aligned}$$

for all s , $0 \leq s < \infty$.

5.4 The Monotonicity of the Power
Function of T_0^2

Let $\Lambda:(m \times m)$ and $\Lambda^*:(m \times m)$ be diagonal matrices with diagonal elements $\lambda_1, \dots, \lambda_m$ and $\lambda_1^*, \dots, \lambda_m^*$ respectively, and suppose that

$$\lambda_i \geq \lambda_i^* \quad \text{for some } i,$$

and

$$(5.4.1)$$

$$\lambda_j = \lambda_j^* \quad \text{for } j \neq i.$$

We shall demonstrate that

$$\beta(\Lambda; m, v_1, v_2; t) \geq \beta(\Lambda^*; m, v_1, v_2; t) \quad (5.4.2)$$

for all $t \geq 0$.

With the representation of T given in (3.2.34) and an obvious notational convention, we have

$$T = \sum_{i=1}^m \frac{u_i^2}{d_i} \quad (5.4.3)$$

and

$$T^* = \sum \frac{u_i^{*2}}{d_i}, \quad (5.4.4)$$

in which (see equations (3.2.28) and (3.3.6))

$$u_i^2 | K \wedge \chi_{v_1}^2(\gamma_i) \quad (5.4.5)$$

and

$$u_i^{*2} | K \wedge \chi_{v_1}^2(\gamma_i^*), \quad (5.4.6)$$

where (see equation (3.3.5))

$$\gamma_i = \sum_{j=1}^m k_{ij}^2 \lambda_j, \quad (5.4.7)$$

and

$$\gamma_i^* = \sum_{j=1}^m k_{ij}^2 \lambda_j^*. \quad (5.4.8)$$

Recall that conditional on K , $u_1^2, \dots, u_m^2$ are mutually independent as are $u_1^{*2}, \dots, u_m^{*2}$. With

$$\alpha_i = \frac{1}{d_i} : i = 1, \dots, m, \quad (5.4.9)$$

and for fixed $\underline{\alpha}' = (\alpha_1, \dots, \alpha_m)$, consider the function

$$\begin{aligned} Q(\theta) &= \int_0^\infty (P(T > t | \underline{\alpha}, K, \Lambda) - P(T^* > t | \underline{\alpha}, K, \Lambda^*)) e^{-\theta t} dt \\ &= \int_0^\infty \left[P \left(\sum_{i=1}^m \alpha_i u_i^2 > t | \underline{\alpha}, K, \Lambda \right) \right. \\ &\quad \left. - P \left(\sum \alpha_i u_i^{*2} > t | \underline{\alpha}, K, \Lambda^* \right) \right] e^{-\theta t} dt. \end{aligned} \quad (5.4.10)$$

Equation (5.4.2) is proven if we can show that $Q(\theta)$ is completely monotone (θ) for all $\alpha_i > 0$: $i = 1, \dots, m$.

Direct integration of (5.4.10) yields

$$Q(\theta) = \frac{1}{\theta} \left| E \left[\exp \left[-\theta \sum_{i=1}^m \alpha_i u_i^{*2} | \underline{\alpha}, K \right] \right] \right|$$

$$= E \left[\exp \left[-\theta \sum_{i=1}^m \alpha_i u_i^2 \mid \underline{\alpha}, K \right] \right]. \quad (5.4.11)$$

Using (5.4.5) and (5.4.6) and taking expectations conditional on K , we have

$$\begin{aligned} Q(\theta) &= \frac{1}{\theta} R(\theta) E_K \left[\exp \left[-\sum_{i=1}^m \beta_i(\theta) \gamma_i^* \right] \right. \\ &\quad \left. - \exp \left[-\sum_{i=1}^m \beta_i(\theta) \gamma_i \right] \mid \underline{\alpha}, K \right] \end{aligned} \quad (5.4.12)$$

where

$$R(\theta) = \prod_{i=1}^m (1 + 2\theta\alpha_i)^{-\nu_1/2} \quad (5.4.13)$$

and

$$\beta_i(\theta) = \frac{\theta\alpha_i}{1 + 2\theta\alpha_i}. \quad (5.4.14)$$

For $r = 1, \dots, m$ let

$$\begin{aligned} q_r(\theta) &= E_K \left[\exp \left[-\left(\sum_{i=1}^r \beta_i(\theta) \gamma_i^* - \sum_{i=r+1}^m \beta_i(\theta) \gamma_i \right) \right] \right. \\ &\quad \left. - \exp \left[-\left(\sum_{i=1}^{r-1} \beta_i(\theta) \gamma_i^* - \sum_{i=r}^m \beta_i(\theta) \gamma_i \right) \mid \underline{\alpha}, K \right] \right]. \end{aligned} \quad (5.4.15)$$

Then we may write

$$Q(\theta) = \frac{1}{\theta} R(\theta) \sum_{r=1}^m q_r(\theta). \quad (5.4.16)$$

From equations (5.4.1), (5.4.7) and (5.4.8),

$$\begin{aligned} P(\gamma_r > s | \gamma_1^*, \dots, \gamma_{r-1}^*, \gamma_{r+1}, \dots, \gamma_m) \\ \geq P(\gamma_r^* > s | \gamma_1^*, \dots, \gamma_{r-1}^*, \gamma_{r+1}, \dots, \gamma_m) \end{aligned} \quad (5.4.17)$$

for all $s > 0$ and $r = 1, \dots, m$. Therefore (see equation (5.3.6))

$$\begin{aligned} \psi_r(\delta) &= \int_0^\infty \left[P(\gamma > s | \gamma_1^*, \dots, \gamma_{r-1}^*, \gamma_{r+1}, \dots, \gamma_m) \right. \\ &\quad \left. - P(\gamma_r^* > s | \gamma_1^*, \dots, \gamma_{r-1}^*, \gamma_{r+1}, \dots, \gamma_m) \right] e^{-\delta s} ds \\ &= \frac{1}{\delta} \frac{E}{K} \left[e^{-\delta \gamma_r^*} - e^{-\delta \gamma_r} \Big|_{\gamma_1^*, \dots, \gamma_{r-1}^*, \gamma_{r+1}, \dots, \gamma_m} \right] \end{aligned} \quad (5.4.18)$$

is completely monotone (δ) . Since $\beta_i(\theta)$ (see (5.4.14)) is a positive function with a completely monotone derivative, then from (5.3.4),

$$\begin{aligned} \psi_r[\beta_r(\theta)] &= \frac{1}{\beta_r(\theta)} \frac{E}{K} \left[e^{-\beta_r(\theta) \gamma_r^*} \right. \\ &\quad \left. - e^{-\beta_r(\theta) \gamma_r} \Big|_{\gamma_1^*, \dots, \gamma_{r-1}^*, \gamma_{r+1}, \dots, \gamma_m} \right] \end{aligned} \quad (5.4.19)$$

is completely monotone (θ) . This in turn implies that

$$\frac{1}{\beta_r(\theta)} q_r(\theta) = \frac{1 + 2\theta\alpha_r}{\theta\alpha_r} q_r(\theta) \quad (5.4.20)$$

is completely monotone (θ) . Now

$$\frac{1}{\theta} \beta_r(\theta) = \frac{\alpha_r}{1 + 2\theta\alpha_r} \quad (5.4.21)$$

is completely monotone (θ) , and hence, using (5.3.3) and (5.4.20), $\frac{1}{\theta} q_r(\theta)$ is completely monotone (θ) . Direct differentiation shows that $R(\theta)$ [see equation (5.4.13)] is completely monotone (θ) and, using (5.4.16) along with an application of property 2 of completely monotone functions [see equations (5.3.2) and (5.3.3)], implies that $Q(\theta)$ is completely monotone (θ) . Equation (5.3.6) implies that

$$P(T > t | \underline{\alpha}, \Lambda) \geq P(T^* > t | \underline{\alpha}, \Lambda^*) \quad (5.4.22)$$

for all $t \geq 0$. Since (5.4.22) is true for all $\underline{\alpha}$, it is true unconditionally. This demonstrates the result of Das Gupta, *et al.*,

$$\beta(\Lambda; m, v_1, v_2; t) \geq \beta(\Lambda^*; m, v_1, v_2; t) \quad (5.4.23)$$

for all $t \geq 0$.

5.5 Summary

Little is known concerning the behavior of the power function of T_0^2 for general m , v_1 and v_2 , with the

exception that it is an increasing function of each λ_i .

Upper and lower bounds on the integral of the power function of T_0^2 were derived in Sections 5.2 and 5.3.

These are

$$\int_{S(\Lambda; m, v_1, \Xi)} \beta(\Lambda; m, v_1, v_2; t) d\Xi$$

$$\leq A(\Lambda; m, v_1, \Xi) P\left[\chi_{mv_1}^2(\text{tr } \Lambda) / \chi_{mv_1}^2(0) > c\right] \quad (5.5.1)$$

and

$$\int_S^\infty \beta(\Lambda; m, v_1, v_2; t) dt$$

$$\geq \int_S^\infty P\left[\chi_{mv_1}^2(\text{tr } \Lambda) \geq (v_2 - m - 1)t\right] dt, \quad (5.5.2)$$

where in (5.5.1) t and c are chosen such that

$$\beta((0); m, v_1, v_2; t) = P\left[\chi_{mv_1}^2(0) / \chi_{mv_2}^2(0) > c\right]$$

$$= \alpha \quad (5.5.3)$$

for some desired level of significance, α .

Although we have not yet been able to prove this, it would seem most likely that $\beta(\Lambda; m, v_1, v_2; t)$ is an increasing function of v_2 . If this is true, an upper bound can be placed on the power of T_0^2 , since

$$\begin{aligned}
 \beta(\Lambda; m, v_1, v_2; t) &= P(T_0^2 > t | \Lambda) \\
 &= P(v_2 \operatorname{tr} Y' S_2^{-1} Y > t | \Lambda) \\
 &= P(\operatorname{tr} Y' (S_2/v_2)^{-1} Y > t | \Lambda).
 \end{aligned}
 \tag{5.5.4}$$

If the conjecture is true, then

$$\begin{aligned}
 \beta(\Lambda; m, v_1, v_2; t) &\leq \lim_{v_2 \rightarrow \infty} P(\operatorname{tr} Y' (S_2/v_2)^{-1} Y > t | \Lambda) \\
 &= P(\chi_{mv_1}^2(\operatorname{tr} \Lambda) > t'),
 \end{aligned}
 \tag{5.5.5}$$

where t' satisfies

$$\begin{aligned}
 \beta((0); m, v_1, v_2; t) &= P(\chi_{mv_1}^2(0) > t') \\
 &= \alpha.
 \end{aligned}
 \tag{5.5.6}$$

We repeat that this is merely a conjecture, but would seem a fruitful area for future research.

BIBLIOGRAPHY

- [1] Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.
- [2] Bagai, O. P. (1962). Statistics Proposed for Various Tests of Hypotheses and Their Distributions. *Sankhyā Ser. A* 24 409-418.
- [3] Constantine, A. G. (1963). Some Non-Central Distribution Problems in Multivariate Analysis. *Ann. Math. Statist.* 34 1270-1285.
- [4] Constantine, A. G. (1966). The Distribution of Hotelling's Generalized T_o^2 . *Ann. Math. Statist.* 37 215-225.
- [5] Das Gupta, S., Anderson, T. W., and Mudholkar, G. S. (1964). Monotonicity of the Power Functions of Some Tests of the Multivariate Linear Hypothesis. *Ann. Math. Statist.* 35 200-205.
- [6] Davis, A. W. (1968). A System of Linear Differential Equations for the Distribution of Hotelling's Generalized T_o^2 . *Ann. Math. Statist.* 39 815-832.
- [7] Davis, A. W. (1970). Exact Distribution of Hotelling's Generalized T_o^2 . *Biometrika* 57 187-191.
- [8] Deemer, W. L. and Olkin, I. (1951). The Jacobians of Certain Matrix Transformations Useful in Multivariate Analysis. *Biometrika* 38 345-367.
- [9] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications*. 2. Wiley, New York.
- [10] Ghosh, M. N. (1964). On the Admissibility of Some Tests of MANOVA. *Ann. Math. Statist.* 35 789-794.
- [11] Hotelling, H. (1931). The Generalization of Student's Ratio. *Ann. Math. Statist.* 2 360-378.
- [12] Hotelling, H. (1947). Multivariate Quality Control, Illustrated by Air Testing of Sample Bombsights. *Techniques of Statistical Analysis*, 11-184. McGraw-Hill, New York.

- [13] Hotelling, H. (1951). A Generalized T Test and Measure of Multivariate Dispersion. *Proc. Second Berkley Symp. Prob. Statist.* 23-41.
- [14] Hsu, P. L. (1938). Notes on Hotelling's Generalized T. *Ann. Math. Statist.* 9 231-243.
- [15] Hsu, P. L. (1940). On Generalized Analysis of Variance, (I). *Biometrika* 31 221-237.
- [16] Hsu, P. L. (1945). On the Power Function of the E^2 -Test and the T^2 -Test. *Ann. Math. Statist.* 16 278-286.
- [17] Ito, K. (1956). Asymptotic Formulae for the Distribution of Hotelling's Generalized T^2 Statistic. *Ann. Math. Statist.* 27 1091-1105⁰.
- [18] Ito, K. (1960). Asymptotic Formulae for the Distribution of Hotelling's Generalized T^2 Statistic, II. *Ann. Math. Statist.* 31 1148-1153⁰.
- [19] Ito, K. (1962). A Comparison of the Powers of Two Multivariate Analysis of Variance Tests. *Biometrika* 49 455-462.
- [20] James, A. T. (1960). The Distribution of Latent Roots of the Covariance Matrix. *Ann. Math. Statist.* 31 151-158.
- [21] James, A. T. (1961). Zonal Polynomials the Real Positive Definite Symmetric Matrices. *Ann. Math.* 74 456-469.
- [22] James, A. T. (1964). Distributions of Matrix Variates and Latent Root Derived from Normal Samples. *Ann. Math. Statist.* 35 473-501.
- [23] Khatri, C. G. and Pillai, K. C. S. (1965). Some Results on the Noncentral Multivariate Beta Distribution and Moments of Traces of Two Matrices. *Ann. Math. Statist.* 36 1511-1520.
- [24] Khatri, C. G. and Pillai, K. C. S. (1966). On the Moments of the Trace of a Matrix and Approximations to Its Noncentral Distribution. *Ann. Math. Statist.* 37 1312-1318.
- [25] Khatri, C. G. and Pillai, K. C. S. (1968). On the Moments of Elementary Symmetric Functions of the Roots of Two Matrices and Approximations to a Distribution. *Ann. Math. Statist.* 39 1274-1281.

- [26] Lawley, D. N. (1938). A Generalization of Fisher's Z-Test. *Biometrika* 30 180-187.
- [27] Mikhail, N. N. (1965). A Comparison of Tests of the Wilks-Lawley Hypothesis in Multivariate Analysis. *Biometrika* 52 149-156.
- [28] Pearson, E. S. and Hartley, H. O. (1966). *Biometrika Tables for Statisticians, Volume I*. Cambridge University Press, London.
- [29] Pillai, K. C. S. (1954). *On Some Distribution Problems in Multivariate Analysis*. Mimeo. Series No. 88, University of North Carolina, Chapel Hill.
- [30] Pillai, K. C. S. and Jayachandran, K. (1967). Power Comparisons of Tests of Multivariate Hypothesis Based on Four Criteria. *Biometrika* 54 195-210.
- [31] Pillai, K. C. S. and Samson, P. (1959). On Hotelling's Generalization of T^2 . *Biometrika* 46 160-168.
- [32] Saw, J. G. (1964). *Likelihood Ratio Tests of Hypothesis on Multivariate Populations, Volume I: Distribution Theory*. Virginia Polytechnical Institute, Blacksburg.
- [33] Srivastava, J. M. (1968). Power of the Lawley-Hotelling Trace Test in Multivariate Analysis Is a Function of the Population Trace. Abstract. *Ann. Math. Statist.* 39 696.
- [34] Stein, C. (1956). The Admissibility of Hotelling's T^2 -Test. *Ann. Math. Statist.* 27 616-623.
- [35] Wald, A. (1942). On the Power Function of the Analysis of Variance Test. *Ann. Math. Statist.* 13 434-439.
- [36] Wishart, J. (1928). The Generalized Product Moment Distribution in Samples From a Normal Multivariate Population. *Biometrika* 20 52-52.

BIOGRAPHICAL SKETCH

David Timothy Hughes was born November 24, 1943 in Philadelphia, Pennsylvania. He was graduated from Abington High School in June of 1961 and entered Bucknell University the following September. It was there that he met Mr. Paul Benson who was to have a profound influence on his life. Mr. Benson introduced him to statistics, and four years later, after receiving his Bachelor of Science degree with a major in Mathematics, Mr. Hughes enrolled in the Graduate School of the University of Florida. He worked as a teaching assistant in the Department of Statistics and received the degree of Master of Statistics in June of 1967. From that time until the present, he has pursued his work toward the degree of Doctor of Philosophy.

David Timothy Hughes is married to the former Margaret Anne Parrish.

This dissertation was prepared under the direction of the chairman of the candidate's supervisory committee and has been approved by all members of that committee. It was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council, and was approved as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1970

Ernest H. Cox
Dean, College of Arts and
Sciences

Dean, Graduate School

Supervisory Committee:

J. G. Saw
Chairman

P. V. Rao

L. L. Schaeffer

Lynman

Zoran R. Pij-Stojanovic

7 9 46 B²⁷